Vis/NIR spectroscopic measurement of selected fertility parameters of Cuban agricultural soils

Ahmed CHACÓN IZNAGA

Supervisors:
Prof. dr. ir. J. De Baerdemaeker, KU Leuven
Prof. dr. ir. W. Saeys, KU Leuven

Members of the Examination Committee:
Prof. A. Van Loey, KU Leuven, chair
Prof. M. Rodríguez Orozco, UCLV, co-supervisor
Prof. E. Aguila Alcantara, UCLV, co-supervisor
Prof. H. Ramon, KU Leuven
Prof. R. Merckx, KU Leuven
Prof. H. Vandendriessche, KU Leuven
Prof. M. Herrera Suárez, UCLV

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor in Bioscience Engineering

December 2014

Doctoraatsproefschrift nr. 1227 aan de faculteit Bio-ingenieurswetenschappen van de KU Leuven

ISBN 978-90-8826-391-0
D/2014/11.109/68

The research was supported by the VLIR (Vlaamse Interuniversitaire Raad)-UOS Programme for Institutional University Cooperation (IUC)

To my Fiancée

Above all, my heartfelt love and affection goes to my beloved Fiancée Ariany Colás Sánchez, for her personal effort and great patience at all times. For always giving me her invaluable aids throughout these years. There are no words which I can express how grateful I am for your love and support and how very much I love and appreciate you. For which, my mere expression of thanks likewise does not enough.

Acknowledgements

Foremost, I am greatly indebted to my promoter Prof. Dr. ir. Wouter Saeys, for the constantly support of my PhD. study and research, great patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in every stage and writing of this research thesis. He provided me an example and pride in a job well done. All the time he spent, the instructions and opportunities he offered me, inspired me for my future researches on Agronomy and Agricultural engineering topics. I could not have imagined having a better promoter and mentor and I hope you always know how proud I am of you and how grateful I am for you.

I thank the rest of my supervisors, co-supervisors, assessors and thesis committee members - Prof. Josse de Baerdemaeker, Prof. Miguel Rodríguez Orozco, Prof. Edith Aguila Alcantara, Prof. Herman Ramon, Prof. Roel Merckx, Prof. Hilde Vandendriessche and Prof. Seppe Deckers - for reviewing my work and providing constructive feedback.

My sincere and special thanks go to Prof. Miguel Herrera Suárez who, advised, inspired and motivated this research topic. His encouragement, insightful comments and friendship, were for an inestimable value in every moment.

I specially thank to my Professors Dr. Joaquín Machado de Armas, Dr. Pedro Cairo Cairo and Dr. Cristóbal Ríos Albuerne for encouraging my research and for allowing me to grow as a research scientist. Your advices on research as well as on my career have been invaluable.

Deep appreciation is also dedicated to Iroel Sánchez Rodríguez, Bart Kemps, Soner Akpinar, Laetitia Six and Peter Salaets for leading some of my working tasks on different stages of this PhD. research.

Special gratitude goes to my fellow lab mate in KU Leuven Chyngyz Erkinbaev for his kindness, friendship, support, stimulating discussions, and for all the fun we have had in the last three years. Also, I thank to my other lab colleagues Sandeep Sharma and Elvis López Bravo for their selfless dedication to my research work before deadline. I am very grateful to them for the welcoming academic and social environment.

I remain grateful to my colleagues from the UCLV, MSc. Amilcar Barreda Valdés and from the Meteorological Centre of Villa Clara province, MSc. Ismabel Domínguez Hurtado for their particular friendship, which has been invaluable on both an academic and a personal level.

I am most grateful to the analytical chemistry laboratory technicians Oralia, Periquín, Carrazana, Robert, Mabel, Taimí from the Territorial Station for Sugar Cane Research (ETICA, in its Spanish acronym) and the Agricultural Research Centre (CIAP, in its Spanish acronym) belonging to the UCLV of Las Villas. I thank them for the reference soil chemical analyses with conventional methods and for their friendship.

I would like to acknowledge the financial, academic and technical support of the the VLIR-UOS Programme for Institutional University Cooperation (IUC) that provided the necessary financial support for this research.

I express gratitude to the Faculty of Agricultural and Animal Sciences from the UCLV, specially the Department of Agronomy and the coordinator of the Doctoral Programme Prof. Osvaldo Fernández Martínez, for the invaluable support and assistance since the beginning of my PhD. research.

Last but by no means least; I would deeply like to express my respect and appreciation for my mother Laura Iznaga Zulueta, my late father Juan Chacón Urrutia for giving birth to me at the first place and supporting me spiritually throughout my life.

I also thank to my lovely fiancée Ariany Colás, my brother Amílkar Chacón, my niece María Fernanda Chacón, my mother in law Caridad Sánchez, my brother in law Adrián Colás, my late uncle in law Antolín Sánchez and my aunt in law Esperanza Sánchez.

Ahmed Chacón Iznaga
Leuven, December 2014

Abstract

Visible (Vis) and Near Infrared Reflectance (NIR) spectroscopy has been recognised as a rapid-response analytical tool to predict soil fertility parameters. In this respect, the conventional methods frequently used in Cuba to determine some fertility parameters important for sugarcane production, such as organic matter (OM), available phosphorus (P) and potassium (K_2O), are difficult, costly, and time-consuming for practical use. Also, it has been demonstrated that there is a gap at present between the applied fertiliser rates based on present recommendations and the real requirements of the crop, which implies a significant loss in sugarcane production. Therefore, this study was undertaken to test the accuracy of calibration models of soil fertility parameters obtained in laboratory conditions from air dried samples of different fields of Villa Clara province. The correlation coefficients of P and K_2O with the OM were taken into account to achieve this purpose. The parameters P and K_2O, which are not spectrally active in the Vis/NIR range are better predicted when are highly correlated with OM. Also, the wavelength intervals to simplify this methodology were selected.

The soil samples were collected from Cambisol and Vertisol groups of 10 Agroindustrial complexes from the plough layer (0 – 20 cm). The samples were split into two datasets, one for calibration on the landscape context and the other for validation on an independent field. For samples used in calibration set a sampling scheme proposed by the Fertilisers and Amendments Recommendations Service (SERFE, in its Spanish acronym) from Cuba was used. The soil samples for the independent validation sets were selected from two municipalities from Villa Clara province (Santa Clara and Sagüa la Grande, for Cambisol and Vertisol respectively). The reflectance spectra were acquired in laboratory conditions for all the soil samples by using a portable Vis/NIR spectrophotometer in reflectance mode in the wavelength range of 399-1697 nm.

The regression models were built in Matlab 7.9 (Mathworks, 2009) by means of Partial Least Squares (PLS) regression, Locally Weighted Regression (LWR) and Support Vector Machine (SVM). These regression methods enabled the relating of near infrared

reflectance spectra to measured values of OM, K_2O and P in the soil. The pre-processing method included Log (1/R), Smoothing and Mean Centre. This method was used while developing the regression models for handling the possible interferences which do not carry chemical information. The PLS, LWR and SVM analyses were evaluated in Venetian blinds cross validation to optimize the model complexity for reliable prediction of these soil fertility parameters.

Results indicate a significant spatial variability of all essential soil fertility parameters at landscape level and at field scale; however there was less variation in the OM content than in the P and K_2O content. Also, all the regression models (PLS, LWR and SVM) provided good correlations between soil spectra and OM. The best accuracy corresponded to the nonlinear regression models for Cambisol and Vertisol at landscape level and within a field scale. For the prediction of the average soil fertility parameters at landscape level on Cambisol soil the best results were obtained for OM ($0.90 \leq R^2 \leq 0.93$; $0.12 \leq RMSEP \leq 0.14$), followed by K_2O ($0.77 \leq R^2 \leq 0.79$; $3.47 \leq RMSEP \leq 3.62$), Olsen P ($0.69 \leq R^2 \leq 0.81$; $0.27 \leq RMSEP \leq 0.35$) and Oniani P ($0.64 \leq R^2 \leq 0.65$; $3.31 \leq RMSEP \leq 3.54$). The prediction accuracy at landscape level on Vertisol soil were similar for OM ($0.81 \leq R^2 \leq 0.87$; $0.16 \leq RMSEP \leq 0.22$) and for K_2O ($0.83 \leq R^2 \leq 0.87$; $2.09 \leq RMSEP \leq 2.40$). However, these results were better than for Olsen P ($0.76 \leq R^2 \leq 0.80$; $0.55 \leq RMSEP \leq 0.67$) and Oniani P ($0.74 \leq R^2 \leq 0.81$; $2.86 \leq RMSEP \leq 3.10$). The results for the prediction of variation in soil fertility parameters within a field on Cambisol soil showed successful correlations between soil spectra and OM ($R^2=0.92$; $RMSEP=0.14$). Also, in this type of soil the non-linear regression models gave the best results for K_2O ($0.61 \leq R^2 \leq 0.63$; $5.13 \leq RMSEP \leq 5.88$), Olsen P ($0.68 \leq R^2 \leq 0.83$; $0.27 \leq RMSEP \leq 0.34$) and Oniani P ($0.70 \leq R^2 \leq 0.72$; $2.32 \leq RMSEP \leq 2.52$). Within a field on Vertisol soil the results were lower than those obtained in Cambisol. The results obtained for OM ($0.79 \leq R^2 \leq 0.80$; $0.21 \leq RMSEP \leq 0.24$) were higher than K_2O ($0.60 \leq R^2 \leq 0.61$; $3.03 \leq RMSEP \leq 3.37$), Olsen P ($0.51 \leq R^2 \leq 0.58$; $0.59 \leq RMSEP \leq 0.78$) and Oniani P ($0.56 \leq R^2 \leq 0.58$; $2.91 \leq RMSEP \leq 4.23$). These results promoted the basic knowledge for applying a strategy of precision fertilisation in Villa Clara province.

Samenvatting

Vis/NIR-spectroscopie werd geïdentificeerd als een techniek met potentieel voor snelle meting van de belangrijkste bodemvruchtbaarheidsparameters. De in Cuba gangbare methode voor het bepalen van de belangrijkste vruchtbaarheidsparameters voor suikerrietproductie, zoals organische stof (OM), beschikbare fosfor (P) en kalium (K_2O), zijn te moeilijk, duur en tijdrovend voor praktisch gebruik. Bovendien werd aangetoond dat er op dit moment een kloof is tussen de toegepaste meststofdosis op basis van deze aanbevelingen en de werkelijke behoeften van het gewas. Dit resulteert soms in een significant verlies in suikerrietproductie. Om hierop een antwoord te bieden, werd in deze studie de nauwkeurigheid onderzocht van de calibratiemodellen voor voorspelling van de bodemvruchtbaarheidsparameters op basis van Vis/NIR-spectra opgemeten in laboratoriumomstandigheden op luchtgedroogde monsters van verschillende gebieden uit de Cubaanse provincie Villa Clara. Hierbij werd rekening gehouden met de correlatie tussen P en K_2O met OM. Voor de parameters P en K_2O die in het Vis/NIR-bereik niet spectraal actief zijn, werd een betere voorspelling bekomen wanneer ze sterk gecorreleerd waren met het gehalte OM. Daarnaast werden ook de belangrijkste golflengte-intervallen geselecteerd om tot eenvoudigere modellen te komen.

De bodemstalen werden bemonsterd in de ploegzone (0-20cm) op velden van 10 agro-industriële complexen met Cambisol- en Vertisolbodems. De monsters werden gesplitst in twee datasets, een calibratieset op landschapsniveau en een validatieset met stalen van eenzelfde onafhankelijk veld. De monsters werden verzameld en geanalyseerd volgens de procedures van de Dienst voor Adviezen rond Meststoffen en Bodemverbeteraars (SERFE). De bodemmonsters voor de onafhankelijke validatieset werden geselecteerd uit twee gemeenten van de provincie Villa Clara (Santa Clara en Sagua la Grande, respectievelijk voor Cambisol en Vertisol). De reflectiespectra werden onder laboratoriumomstandigheden opgemeten voor alle bodemmonsters met behulp van een draagbare Vis/NIR-spectrofotometer in reflectiemodus in het golflengtegebied van 399-1697 nm. De regressiemodellen tussen de opgemeten NIR-reflectiespectra en

de OM, K_2O en P gehaltes in de bodem werden in Matlab 7.9 (Mathworks, 2009) gebouwd door middel van Partiële Kleinste Kwadratenregressie (PLS), Lokaal Gewogen Regressie (LWR) en Support Vector Machines (SVM). De spectra werden voorbehandeld door spectrale transformatie (Log(1/R)), filtering en centeren rond het gemiddelde. Deze spectrale voorbehandeling werd toegepast tijdens de ontwikkeling van de regressiemodellen om het effect van de mogelijke interferenties die geen chemische informatie dragen te verwijderen. De PLS, LWR en SVM-analyses werden geëvalueerd in jaloezieën-kruisvalidatie om de complexiteit van het model voor een betrouwbare voorspelling van de bodemvruchtbaarheidsparameters te optimaliseren.

De bekomen resultaten toonden een aanzienlijke ruimtelijke variabiliteit van alle essentiële bodemvruchtbaarheidsparameters op landschapsniveau en op veldschaal. De variatie in het OM-gehalte was relatief gezien kleiner dan deze in het P en K_2O-gehalte. Alle regressiemodellen (PLS, LWR en SVM) gaven een goede correlatie tussen de bodem spectra en het OM-gehalte. De beste nauwkeurigheid werd voor zowel Cambisol als Vertisol op landschapsniveau en binnen een veld bekomen met de niet-lineaire regressiemodellen. Voor de voorspelling op landschapsniveau van de gemiddelde bodemvruchtbaarheidsparameters in Cambisol-bodems werden de beste resultaten verkregen voor OM ($0.90 \leq R^2 \leq 0.93$; $0.12 \leq RMSEP \leq 0.14$), gevolgd door K_2O ($0.77 \leq R^2 \leq 0.79$; $3.47 \leq RMSEP \leq 3,62$), Olsen P ($0.69 \leq R^2 \leq 0.81$; $0.27 \leq RMSEP \leq 0.35$) en Oniani P ($0.64 \leq R2 \leq 0.65$; $3.31 \leq RMSEP \leq 3.54$). De voorspellingsnauwkeurigheid op landschapsniveau voor Vertisol-bodems was vergelijkbaar voor OM ($0.81 \leq R^2 \leq 0.87$; $0.16 \leq RMSEP \leq 0.22$) en voor K_2O ($0.83 \leq R^2 \leq 0.87$; $2.09 \leq RMSEP \leq 2.40$). Voor Olsen P ($0.76 \leq R^2 \leq 0.80$, $0.55 \leq RMSEP \leq 0.67$) en Oniani P ($0.74 \leq R^2 \leq 0.81$; $2.86 \leq RMSEP \leq 3.10$) waren de resultaten zelfs beter. De resultaten voor de voorspelling van de variatie in bodemvruchtbaarheidparameters binnen een veld met Cambisol-bodem toonden hoge correlaties tussen de bodemspectra en het OM-gehalte ($R^2 = 0.92$; $RMSEP = 0.14$). In dit bodemtype gaven de niet-lineaire regressiemodellen ook de beste resultaten voor K_2O ($0.61 \leq R^2 \leq 0.63$, $5.13 \leq RMSEP \leq 5.88$), Olsen P ($0.68 \leq R^2 \leq 0.83$, $0.27 \leq RMSEP \leq 0.34$) en Oniani P ($0.70 \leq R2 \leq 0.72$; $2.32 \leq RMSEP \leq 2.52$). De voorspelling van de variatie binnen een veld met Vertisol-bodem was minder goed dan voor het veld met Cambisol-bodem. De resultaten verkregen voor OM ($0.79 \leq R^2 \leq 0.80$; $0.21 \leq RMSEP \leq 0.24$) waren beter dan deze voor K_2O ($0.60 \leq R^2 \leq 0.61$; $3.03 \leq RMSEP \leq 3.37$), Olsen P ($0.51 \leq R^2 \leq 0.58$; $0.59 \leq RMSEP \leq 0.78$) en Oniani P ($0.56 \leq R^2 \leq 0.58$; $2.91 \leq RMSEP \leq 4.23$). Deze resultaten bevorderen het gebruik van deze technologie voor precisiebemesting in de Cubaanse Villa Clara provincie.

Resumen

La espectroscopia en el Visible (Vis) e Infrarrojo cercano (NIR) se reconoce como una herramienta analítica de respuesta rápida, para la determinación de parámetros de fertilidad del suelo. Respecto a lo anterior, los métodos convencionales frecuentemente usados en Cuba para determinar algunos parámetros de fertilidad importantes para la producción de caña de azúcar, tales como materia orgánica, fósforo y potasio disponible, incluyen procedimientos que son a menudo demasiado difíciles, costosos, y consumen mucho tiempo para un uso práctico. También se ha demostrado que en el presente existe una diferencia entre las dosis de fertilizante aplicadas en base a las recomendaciones actuales y las necesidades reales del cultivo, lo cual implica pérdidas significativas en la producción de la caña de azúcar. Por consiguiente, el objetivo de esta investigación estuvo dirigido a comprobar la precisión de modelos de calibración de parámetros de fertilidad de suelo, obtenidos en condiciones de laboratorio a partir de muestras de suelo secadas al aire y colectadas en diferentes campos de la provincia de Villa Clara. Los parámetros de fertilidad P y K_2O, que no son espectralmente activos en el rango Vis/NIR, alcanzan una mejor predicción cuando presentan una alta correlación con la materia orgánica. Además, los intervalos de longitud de onda para simplificar esta metodología fueron seleccionados.

Las muestras de suelo se colectaron en los tipos de suelo Cambisol y Vertisol a una profundidad de 0 - 20 cm. Los suelos se localizan en 10 complejos Agroindustriales de la provincia. Las muestras se dividieron en dos bases de datos, una para la calibración en el contexto de toda la provincia y el otro para la validación en un campo independiente. Para las muestras utilizadas en la calibración, se empleó el esquema de muestreo propuesto por el Servicio de Recomendaciones de Fertilizantes y Enmiendas (SERFE) de Cuba. Las muestras de suelo para las bases de datos empleadas en la validadación en campos independientes, fueron seleccionadas de dos municipios de Villa Clara (Santa Clara y Sagüa la Grande, para Cambisol y Vertisol respectivamente). Los espectros de reflectancia de todas las muestras de suelo fueron obtenidos en condiciones de

laboratorio, mediante la utilización de un espectrofotómetro Vis/NIR portátil en el modo de reflectancia, ajustándose un rango de longitud de onda entre 399 – 1697 nm.

Los modelos de regresión se desarrollaron en Matlab 7.9 (Mathworks, 2009) mediante las técnicas de regresión de mínimos cuadrados parciales (PLS), regresión pesada local (LWR) y máquina de vectores de soporte (SVM). Estas técnicas posibilitaron obtener la relación entre los espectros de reflectancia de infrarrojo cercano y los valores de OM, K_2O y P determinados por métodos convencionales. El método de pre-procesamiento de datos espectrales utilizado incluyó *Log (1/R)*, *Centrado medio* y *Suavizado*. El pre-procesamiento se realizó con el fin de remover los datos irrelevantes que interfieren en la información química. Los modelos de regresión PLS, LWR y SVM fueron evaluados con una técnica de validación cruzada (Persianas venecianas), para optimizar la complejidad del modelo y obtener una predicción fiable de los parámetros de fertilidad del suelo.

Los resultados indicaron que existe una variabilidad espacial significativa de todos los parámetros de fertilidad de suelo estudiados, tanto en el contexto de los suelos de diferentes campos de la provincia, como dentro de un campo específico, sin embargo el contenido de materia orgánica mostró menor variabilidad que el contenido de fósforo y potasio. Además, todos los modelos de regresión (PLS, LWR y SVM) mostraron buenas correlaciones entre el espectro de suelo y la materia orgánica. Las mejores precisiones se obtuvieron con los modelos de regresión no lineales, tanto en Cambisol, como en Vertisol, tanto en el contexto de toda la provincia, como dentro de un campo en específico. En la predicción de los parámetros de fertilidad de suelo Cambisol en el contexto de la provincia, los mejores resultados fueron obtenidos en la materia orgánica ($0.90 \leq R^2 \leq 0.93$; $0.12 \leq RMSEP \leq 0.14$), seguido por el potasio ($0.77 \leq R^2 \leq 0.79$; $3.47 \leq RMSEP \leq 3.62$), fósforo Olsen ($0.69 \leq R^2 \leq 0.81$; $0.27 \leq RMSEP \leq 0.35$) y fósforo Oniani ($0.64 \leq R^2 \leq 0.65$; $3.31 \leq RMSEP \leq 3.54$). En igual contexto los resultados obtenidos en Vertisol fueron similares para la materia orgánica ($0.81 \leq R^2 \leq 0.87$; $0.16 \leq RMSEP \leq 0.22$) y el potasio ($0.83 \leq R^2 \leq 0.87$; $2.09 \leq RMSEP \leq 2.40$). Sin embargo, estos resultados fueron mejores que los obtenidos en el fósforo Olsen ($0.76 \leq R^2 \leq 0.80$; $0.55 \leq RMSEP \leq 0.67$) y fósforo Oniani ($0.74 \leq R^2 \leq 0.81$; $2.86 \leq RMSEP \leq 3.10$). La predicción de los parámetros de fertilidad de suelo Cambisol en un campo específico mostró altas correlaciones entre el espectro de suelo y la materia orgánica ($R^2=0.92$; $RMSEP=0.14$). También, en este tipo de suelo los modelos de regresión no lineal mostraron los mejores resultados en el

potasio ($0.61 \leq R^2 \leq 0.63$; $5.13 \leq RMSEP \leq 5.88$), fósforo Olsen ($0.68 \leq R^2 \leq 0.83$; $0.27 \leq RMSEP \leq 0.34$) y fósforo Oniani ($0.70 \leq R^2 \leq 0.72$; $2.32 \leq RMSEP \leq 2.52$). En un campo específico de Vertisol los resultados fueron menores que los obtenidos en Cambisol. Los resultados de la materia orgánica ($0.79 \leq R^2 \leq 0.80$; $0.21 \leq RMSEP \leq 0.24$), fueron más altos que los de potasio ($0.60 \leq R^2 \leq 0.61$; $3.03 \leq RMSEP \leq 3.37$), fósforo Olsen ($0.51 \leq R^2 \leq 0.58$; $0.59 \leq RMSEP \leq 0.78$) y fósforo Oniani ($0.56 \leq R^2 \leq 0.58$; $2.91 \leq RMSEP \leq 4.23$). Estos resultados proporcionaron el conocimiento básico para la puesta en práctica de una estrategia de fertilización precisa en la provincia de Villa Clara.

List of abbreviations

ANN	Artificial Neural Networks
CAI	Agroindustrial complex (in its Spanish acronym)
$CaSO_4 \cdot 2H_2O$	Gypsum
CCDs	Charge Coupled Devices
CEC	Cation Exchange Capacity
CIAP	Agricultural Research Centre (in its Spanish acronym)
$K_2Cr_2O_7$	Potassium dichromate
CV	Coefficient of Variation
DRS	Diffuse Reflectance Spectroscopy
d.s	dry soil
ETICA	Territorial Station for Sugar Cane Research
EMSC	Extended Multiplicative Signal Correction
EPO	External Projection Orthogonalisation
FWR	Full Wavelength Range
GLSW	Generalized Least Squares Weighting
H_2SO_4	Sulfuric acid
InGaAs	Indium gallium arsenide
K_2O	Potassium content determined by the Oniani method
KH_2PO_4	Monopotassium phosphate
LEDs	Light-emitting diodes
LS-SVM	Least Squares Support Vector Machines
LWR	Locally Weighted Regression
masl	metres above sea level
Max	Maximum observed value
mid-IR	Mid Infrared
Min	Minimum observed value
MSC	Multiplicative Scatter Correction
$NaHCO_3$	Sodium hydrogen carbonate
NIR	Near Infrared Reflectance
OC	Organic Carbon
Olsen P	Phosphorus content determined by the Olsen method
OM	Organic Matter
Oniani P	Phosphorus content determined by the Oniani method
OSC	Orthogonal Signal Correction

PbS	Lead sulfide
P_2O_5	Phosphorus pentoxide
PCA	Principal Component Analyses
PLS	Partial Least Squares
PCR	Principal Component Regression
PSI	Sorption index for phosphorus
RER	Ratio of Error Range
RMSEC	Root Mean Square Error of Calibration
RMSECV	Root Mean Square Error of Cross Validation
RMSEP	Root Mean Square Error of Prediction
RPD	Ratio of Performance Deviation
SD	Standard Deviation
SEP	Standard Error of Prediction
Si	Silicon
SNV	Standard Normal Variate
SVM	Support Vector Machine
SWR	Short Wavelength Range
TC	Total Carbon
UV	Ultra violet
VIS	Visible reflectance

List of symbols

A	Absorbance
C	Concentration of the absorbing species, mol L^{-1}
K	Absorptivity
I	Light intensity
I_0	Intensity of incident light
i	Sample number
L	Pathlength
M	Molarity: refers to the number of molecules of a substance in a solution
N	Size of the test set
R	Reflectance
y	Reference value
$\hat{y}$	Prediction value
r	Coefficient of correlation
R^2	Coefficient of determination

List of figures

List of tables

Table of contents

Chapter 1

General introduction

1.1. Introduction

Agriculture in Cuba has had a complex history of difficulties and extremes. When the current government came to power in January 1959, 75% of Cuba's agricultural land was owned by foreign companies and individuals. The main crop was sugarcane, which was sold to the United States (US) and Great Britain (Hunt, 2008). Since the beginning of the Cuban nation the sugarcane crop and the sugar production have been supporting the economy. Both features are considerably linked to the social development, culture and traditions of Cuba. Before 1959 a large sugar quotum was given to Cuba by North America which paid above world prices in order to support US industry. After the Revolution the new Cuban government adopted a series of land reforms which finally resulted in the confiscation of almost all private property in favor of the establishment of large Soviet-style State farms whose creation arose from the notion that the State was the central force and that heavy mechanization would improve the dignity of human labor. However, the end result was both lost production and 'lost' workers, for the relegation of peasant farms to non-preferred status meant their production failed to develop and disaffected agricultural producers and laborers migrated en masse from rural to urban areas (Hunt, 2008).

Between 1976 and 1990, sugar represented three fourths of the value of Cuba's merchandise exports and 20% of the economy's total investments. By the early 1990's, the sugar industry held 30% of the nation's industrial machinery and bought 20% of the production of the steel and machinery sector, 13% of the production of the ministry of basic industries, and 8% of light industry output (Peters, 2003). The US trade embargo of Cuba, plus the collapse of the island's Soviet market, has meant that the country has found it virtually impossible to import the chemicals and machinery necessary to practice modern, intensive agriculture (Warwick, 2000).

It left Cuba with a large, mechanized sugar industry designed to fit a customer that no longer existed, and dependent on sources of tractors, fuel, fertilisers, and other supplies that also ceased to exist. Repeated in many sectors, this experience plunged Cuba into a severe economic crisis (Peters, 2003).

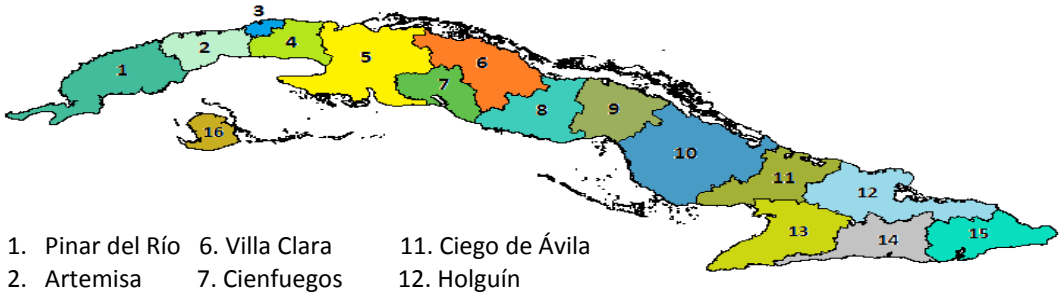
Therefore, Cuba's sugarcane production sharply declined subsequently, from 8.4 million tons in 1990 to 4.2 million only three years later. A blatant lack of efficiency, a series of droughts and hurricanes, as well as an economic crisis led to a fall in average annual production to a mere 3.7 million tons from 1994 to 2003. In 2002 the government, in despair, severely downsized the industry, closing over half of Cuba's 156 Agroindustrial Complexes (CAI, in its Spanish acronym) in what was called the "Alvaro Reynoso Task". As a result, production continued to shrink. By 2007-2008, the Cuban zafra amounted to a mere 1.5 million tons. As a result, sugarcane production at that period was only about one-eighth of the peak reached in 1990 (Alonso-Pippo, 2008; Fischer et al., 2008).

The economic effects of the disappearance of the socialist markets at the decade of the 90 impacted on the use of mineral fertilisers in sugarcane production. Therefore, their use diminished abruptly to 30% (Ruiz, 2004). This type of fertiliser is essential to maintain the soil fertility and to increase the production in the agriculture.

At present, the necessary inputs for the production of fertilisers are scarce and the prices are increased. Thus, for obtaining high efficiency productivity in sugarcane becomes more and more necessary, especially under the natural conditions of the tropics, such as high temperatures, abundant rains and soil characteristics. These factors can increase fertiliser losses by different ways, causing low nutrient use efficiency. Consequently, in order to reduce the expenses for imported inputs, Cuba is forced to look for alternatives to make an appropriate fertilisation. These alternatives have to be aimed at the efficient use of the nutrients that plants need, at the improvement of the soil structural properties and consequently at the increment in crop yields. A suitable balance among the availability of the soil nutrients and the quantities given through the fertilisation process, is essential to rationalize the use of the inputs and to reach economic advantages in the agricultural production.

1.2. Cuban soils and soil fertility overview

The Republic of Cuba is divided in 15 provinces and the special municipality *Isla de la Juventud* (Figure 1-1). It extends about 1 225 km from *Cabo de San Antonio* to *Punta de Maisí* the western and eastern extremities respectively and commands the two entrances to Mexico's Gulf, which are the Florida's Straits and the Yucatan's Channel. The average width is about 80 km, with extremes ranging from 35 to 251 km. The total area is 114 524 km^2 including the area of other islands (FIU, 2012).



1. Pinar del Río
2. Artemisa
3. La Habana
4. Mayabeque
5. Matanzas
6. Villa Clara
7. Cienfuegos
8. Santi Spiritus
9. Ciego de Ávila
10. Camagüey
11. Ciego de Ávila
12. Holguín
13. Granma
14. Santiago de Cuba
15. Guantánamo
16. Isla de la Juventud

Figure 1-1. The Republic of Cuba
Source: Meteorological Centre from Villa Clara province (2011).

Groups of mountains and hills cover about one-fourth of the island of Cuba. The most rugged range is the *Sierra Maestra*, which stretches approximately 240 km along the southeastern coast and reaches the island's highest elevations (1974 m) at *Turquino Peak* and 1730 m at *Bayamesa Peak*. Near the centre of the island are the Santa Clara Highlands, the *Sierra de Escambray* (*Guamuhaya*), and the *Sierra de Trinidad*. The *Cordillera de Guaniguanico* in the far west stretches from southwest to northeast for 180 km and comprises the *Sierra de los Órganos* and the *Sierra del Rosario*, the latter attaining 692 m at *Guajaibón Peak*. Much of central-western Cuba is punctuated by spectacularly shaped, vegetation-clad hillocks called *mogotes*. Serpentine highlands distinguish northern and central Artemisa, Mayabeque, La Habana and Matanzas provinces, as well as the central parts of Camagüey and Las Tunas (Knight and Levinson, 2012). The remainder surface is consisting of flat or rolling land.

The climate is subtropical with a mean annual temperature of 25 ^{0}C. Extremes of heat and relative humidity, which average 27.2 ^{0}C and 80%, respectively, during the summer season, are tempered by the prevailing northeastern trade winds. The annual rainfall averages about 1320 mm. More than 60% of the rain falls during the wet season, which extends from May to October. This nation lies in a region occasionally traversed by violent tropical hurricanes during August, September and November (FIU, 2012).

Knight and Levinson (2012) reported that the complicated Cuban topography and geology have produced at least 13 distinct groups of soils, the majority of which are fertile and cultivated throughout the year. Zhang and Raun (2006) explained that soil fertility is the ability of a soil to provide essential plant nutrients in adequate amounts and proper proportions to sustain plant growth. In this sense, highly fertile red limestone soil extends from the west of Havana to near Cienfuegos on the southern coast. Also, extensive areas of fertile soils lie in the western Camagüey province and north of Cienfuegos between the *Sierra de Sancti Spíritus* and the Caribbean coast, providing the basis for Cuba's main agricultural output. Camagüey province and the Guantánamo basin have some arable land, although of lower fertility. Areas of sandy soil in Pinar del Río, Villa Clara, and portions of Ciego de Ávila and Camagüey provinces cannot hold moisture and are marginally fertile, as are the soils of the mangrove-dotted coastal swamps and cays. Therefore, distinct soil groups can be found in various areas of Cuba (Figure 1-2).

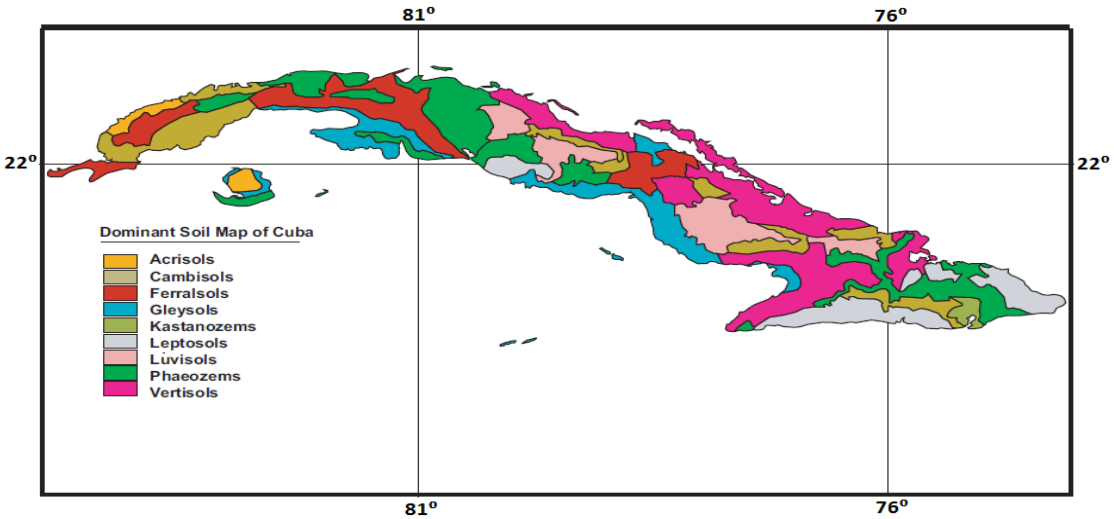


Figure 1-2. Dominant soil map of Cuba

Source: Soil and Terrain Database for Latin America and the Caribbean FAO, ISRIC, UNEP (1998).

The following is a brief description of these soils:

Acrisols are clay-rich, and are associated with humid, tropical climates. The acrisols low fertility and toxic amounts of aluminum pose limitations to its agricultural use, favoring in many places its use for silviculture, low intensity pasture and protected areas. Crops that can be successfully cultivated, if climate allows, include sugar cane.

Cambisols are soils having a cambic horizon. They have mollic overlying subsoil with a base saturation randomly distributed within 100 cm from the soil surface. The horizon differentiation is weak. Most of these soils make good agricultural land and are intensively used. Cambisols in temperate climates are among the most productive soils on earth.

Ferralsols are usually deep soils because both the intensity and the duration of weathering have been considerable. Most ferralsols are strongly coloured. Iron oxides also contribute to the aggregation of clay and silt which creates porosity. Air and water can usually circulate freely though ferralsols; rainfall acceptance by these profiles is usually faster than by most other soils of comparable texture; it also leaches quite rapidly to deeper layers that are beyond the reach of the common rooting of most cultivated crops.

Gleysols are wetland soils (hydric soils) that, unless drained, are saturated with groundwater for long enough periods to develop a characteristic gleyic colour pattern. They exhibit a greenish-blue-grey soil colour due to anoxic wetland conditions. Gley soils may be sticky and hard to work, especially where the gleying is caused by surface water, held up on a slowly permeable layer.

Kastanozems are rich in humus, and originally covered with early maturing native grasslands vegetation, which produces a characteristic brown surface layer in the first metre in depth. These soils have a relative high level of available calcium ions bound to soil particles and can have a petrocalcic horizon between 25 and 100 cm thick.

Leptosols are soils with a very shallow profile depth (indicating little influence of soil-forming processes), and they often contain large amounts of gravel. They typically remain under natural vegetation, being especially susceptible to erosion, desiccation, or water logging, depending on climate and topography.

Luvisols are moderately developed soils from temperate regions. Luvisols have an argillic horizon with a base saturation of at least 50% in the lower part of the B horizon at less than 125 cm depth. The cation exchange capacity of these soils is equal to or more

than 24 moles per kg of clay throughout. There is no E horizon with abrupt transition to a poorly permeable horizon or signs of redistribution of clay. These soils have an accumulation of clays in its subsurface levels.

Phaeozems are soils without carbonates and sulfates in the deeper horizons. These soils present a well structured mollic epipedon, a high base saturation and a moderate to high organic matter content.

Vertisols are soils in which there is a high content of expansive clay known as montmorillonite that forms deep cracks in drier seasons or years. Alternate shrinking and swelling causes *self-mulching*, where the soil material consistently mixes itself, causing vertisols to have an extremely deep A horizon and no B horizon. When irrigation is available, crops can be grown.

According to Hernández et al. (2005) Cuba is one of the few nations which has its own system for soil classification. This began with the studies developed by Bennett and Allison in 1928, which elaborated the North American system at that time about Soil Series and Families. After the Cuban Revolution the approaches of the genetic-geographical school of soil classification began to be introduced with the assistance of specialists from China, Soviet Union, France and Spain. A research project started in 1992 that brought, as a result, the New Version of Soil Genetic Classification of Cuba, being published in 1999. The different soil types in this system are illustrated in Figure 1-3 and their correlations with other soil classification systems are summarised in Table 1-1.

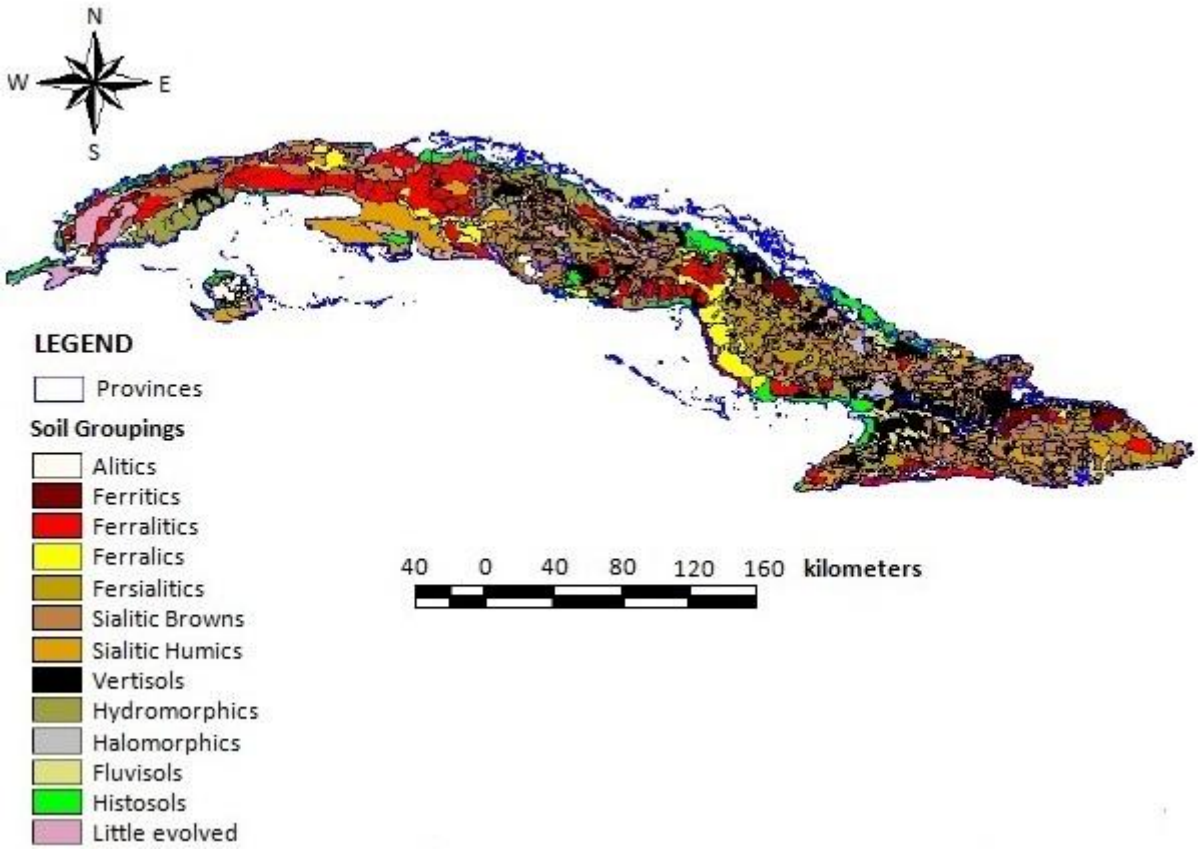


Figure 1-3. Soil map based on the New Version of Soil Genetic Classification of Cuba
Source: Hernández et al. (2001).

Table 1-1

Correlation between different soil groupings classification and the surface represent.

No.	Soil Groupings Classification			Surface	
	Cuba, 1999	**FAO-UNESCO, 1990**	**World Reference Base (WRB)**	**Thousands of hectares**	**%**
I	Alitics	Acrisols	Alisols, Acrisols	606.69	5.67
II	Ferritics	Ferralsols	Ferralsols	207.58	1.94
III	Ferralitics	Ferralsols, Cambisols, Nitisols, Lixisols	Ferralsols, Nitisols, Lixisols	1699.16	15.88
IV	Ferralics	Cambisols	Cambisols, Nitisols, Luvisols	56.71	0.53
V	Fersialitics	Cambisols, Luvisols	Cambisols, Luvisols	13599.97	12.71
VI	Sialitic Browns	Cambisols	Cambisols, Umbrisols	2869.74	26.82
VII	Sialitic Humics	Phaeozems, Leptosols	Phaeozems	743.65	6.95
VIII	Vertisols	Vertisols	Vertisols	881.68	8.24
IX	Hydromorphics	Gleysols	Gleysols	841.02	7.86
X	Halomorphics	-	Solonchaks, Solonetz	87.74	0.82
XI	Fluvisols	Fluvisols	Fluvisols, Cambisols	511.46	4.78
XII	Histosols	Histosols	Histosols	27.82	0.26
XIII	Little evolved	Arenosols, Leptosols	Arenosols, Leptosols	806.78	7.54

The following is a brief description of these soils:

Alitics are soils with ABC profile, with an intense alteration in primary minerals and a saturation degree larger than 50% for exchangeable aluminum, which characterizes the alitic B horizon.

Ferritics are soils with an intense alteration in primary minerals, high iron content and low silica content and variable quantitities of ferruginous nodules.

Ferralitics are soils developed under the ferralitization process, which is characterized by an intense alteration of minerals, formation of clay minerals belonging to the 1:1 type, as well as iron and aluminum oxides and hydroxides.

Ferralics are soils with ABC profile, with red and yellow colours and are characterized for having a ferralic B horizon. These soils are generated from limestones or over schists.

Fersialitics are soils developed under the fersialitization process characterized by the presence of clay minerals and Fe_2O_3 content, with reddish or yellowish colours in the profile or in some horizons (fersialitic).

Sialitic Browns are soils developed under the sialitization process. In general, the free iron content is not greater than 3%.

Sialitic Humics are soils with AC, ACD or rarely A(B)C profiles. The main horizon is humified. Generally, the transition from the A horizon to a lower horizon is abrupt.

Vertisols are mostly clay soils which stick and swell significantly. These soils are not easy for cultivation.

Hydromorphics are soils from areas of water regime. These soils are saturated with water for most parts of the year and hence lack oxygen. The decomposition of organic matter is very slow. As a result, a thick layer of organic matter gets accumulated. This peaty matter gives the soil a dark appearance.

Halomorphics are soils where the total soluble salts are greater than 1% when the soil texture is loamy, are greater than 0.8% when it is clay loam and are greater than 0.6% when it is sandy or sandy loam. When dry, these soils are hard and compacted.

Fluvisols are very young soils with weak horizon differentiation and have mostly AC profiles. These soils are predominantly brown (aerated soils) and/or grey (waterlogged soils) in colour. Their characteristics are dominated by their recent sedimentation and wetness. Most Fluvisols have neutral or near-neutral pH values, which do not impair the availability of nutrients.

Histosols are organic soils formed in water or under water, in which the minimum organic carbon content is 12% (with no clay in the mineral portion) and the maximum is 18 % (with clay in the mineral portion at 60% or above).

Little evolved are soils widely distributed in the mountainous systems and geographic heights of the country, as well as in coastal areas and places where the human activity has caused land degradation.

Soil fertility depends on many factors such as the origin of soil, soil texture and structure, the availability of OM the soil fertility management and the prevailing cropping systems. Gersper et al. (1993) reported that in Cuba the build up and maintenance of soil fertility and productivity is being accomplished with various organic and mineral amendments and biofertilisers (*Azospirillum*, Fitomas-E[1]), produced or mined within the country (locally, where possible) and through rational management utilizing cover crops, green manures, crop successions (intercropping and rotations), and other appropriate technologies.

1.3. Soil fertility management

Sugarcane is the major crop in the Villa Clara province. For growth and development of sugarcane, nitrogen (N), potassium (K) and phosphorous (P) are the most important nutrients, while organic matter (OM) is needed for good soil physical properties. Different management practices such as tillage, fertilisation, manure application, and others can affect these soil fertility parameters. In order to sustain productivity, the nutrients which have been extracted by the harvested crop and which have been lost to the air or leached out to the ground water have to be replenished by adding organic or mineral fertilisers. Calculation of the plant nutrient needs is very important for the rational application of fertilisers; not only for obtaining crop productions which compensate the cost of fertilisers, but also for effective management of the environment. The fertiliser doses should be estimated in order to cover the crop requirements and to maintain or even improve the soil fertility.

Nutrients can be directly added by the application of mineral fertiliser to the soil. However, the addition of fertiliser alone is not enough to retain a sufficient level of soil fertility (van Schöll and Nieuwenhuis, 2004). In this sense, most fertilisers commonly

[1] Fitomas-E is a Cuban-made bio-fertiliser

used in agriculture contain the three basic plant nutrients (N, P, K), for instance, urea (46% N), Triple superphosphate (46% P_2O_5) and Potassium chloride (60% K_2O). Some fertilisers also contain certain micronutrients, such as zinc (Zinc sulphate- 21% Zn) and other metals (Copper sulphate- 24% Cu; Ferrous sulphate- 19% Fe) that are necessary for plant growth. Fertilisers and soil amendments can be derived from virgin raw material, composts and other organic matter, and wastes, as EPA (1997) stated. Fertiliser can be applied in a number of ways such as broadcasting (the fertiliser pellets are spread evenly over the whole field, and then often ploughed or raked into the soil) and row application (the fertiliser is applied in rows).

The application of nutrients (organic / inorganic) in excess may be polluting, especially for the underground water (Sharma, 2005). If improperly managed, these elements can move into surface water through field runoff or leach into ground water. The two most important nutrients that are of greatest concern to water quality (ground water and surface water used as public drinking water supplies) are N and P (EPA, 2010).

Fertiliser prices began to rise steadily in 2002 and increased sharply to historic highs in 2008 due to the combined effects of a number of domestic and global long and shortrun supply and demand factors. From 2007 to 2008, spring nitrogen prices increased by a third, phosphate prices nearly doubled, and potash prices doubled (Huang, 2009). Virtually all fertilisers reached historic highs during this period. The years 2008 and 2009 were very different and 2010 was an average year with no high price spikes, but just an overall gentle price increase. Prices in 2011 particularly for phosphorous have had a large run-up. Dated prices for a fertiliser like 11-52-0 (Monoammonium phosphate) were nearly $880 dollars per ton depending on the market where one was buying in (Horneck, 2011). However, from 2012 to 2013 some major fertiliser products used in agricultural production posted a relatively low decline in prices per ton: urea ($400 to $350) and Diammonium phosphate ($670 to $465).

1.4. SERFE soil analysis system

Cuba, in spite of the serious economic crisis due to the loss of preferential markets and to the low sugar prices on the world market, was able to implement in 1997 the Fertilisers and Amendments Recommendations Service (SERFE in its Spanish acronym). Since 1998 this service has been providing the technical basis to sugarcane producers for optimizing the fertiliser doses to diagnose the mineral fertiliser necessities for this

crop. SERFE is an integrated advisory system for fertiliser recommendations on sugar cane, which includes soil analyses, demonstration plots, advice, annual training courses, developed specialized decision-making software, recommended rates and other actions associated with efficient fertiliser use (FAO, 2003). Applying this system, it is possible to analyse one representative sample per field, but this involves a lot of sample logistics and does not allow determining the variation in soil fertility within a field. In this system, the general fertility analyses commonly include phosphorus, potassium and pH. The conventional methods frequently used to determine these soil fertility parameters are Walkley-Black method (Walkley and Black, 1934) for determining organic matter (OM%) and Oniani method (Oniani et al., 1973) for analysing phosphorus (commonly expressed as mg P_2O_5 100 g^{-1} dry soil) and potassium content (mg K_2O 100 g^{-1} dry soil), while pH is determined by potentiometry.

1.5. Variation in soil fertility: the need for precision fertilisation

According to the considerations of Haneklaus and Schnug (2006), soils are not static, neither homogeneous in space nor time. This influences directly the concentration of the available plant nutrients. In this sense, the uniform application of fertilisers inevitably causes an excess or deficit of nutrients in soil that is neither ecologically nor economically viable. This spatial variability in soil nutrients can be responsible for production differences observed within a field during harvest campaign.

Guimarães et al. (1997) and Bai and Wang (2011) emphasized that understanding the spatial distribution pattern of soil fertility parameters in cultivated areas is important for determining soil constraints to plant nutrition and an appropriate management of soil resources and also for optimizing fertilisation and sustaining yield. However, knowing which soil attributes are most important for a particular local or soil series will not be readily apparent until research agronomists and crop scientists have been able to collect site-specific data for a range of soil attributes over a number of years (Hummel and Birrell, 1998).

Therefore, the soil should be characterized chemically with the purpose of determining the potential fertility before establishing the crop. Thus, before affirming how soil fertility is maintained, the causes of soil fertility decline must be stated and one of the steps for it is to take soil samples from agricultural fields, storing and processing them before performing the analyses by reference or conventional methods. This will bring

objective information about the level of these soil fertility parameters. In theory, the regular soil testing employed by the SERFE system helps to ensure that the correct amount of fertiliser is applied to the soil to satisfy the nutrient requirements of the crop. In that way, it also contributes to avoid any risk to the environment by avoiding over-fertilisation. However, De León et al. (2004) demonstrated that there is a gap at present between the applied fertiliser rates based on present recommendations and the real crop requirements. This implies significant losses in sugarcane production and can affect variability patterns of soil fertility parameters. Like most standard nutrient recommendation strategies the SERFE system is based on the determination of an average fertiliser requirement for the field which is translated into a single rate for the entire field. Using this strategy, some areas of the field receive more than the optimum amount of fertiliser while other areas may not be receiving enough.

For instance, when the growth in an under-fertilised plot with less nitrogen than the normal amount is starting to look paler green than the surrounding growth (Figure 1-4), it could indicate that the crop in this plot is suffering from a lack of nitrogen (Kulmala, 2012).

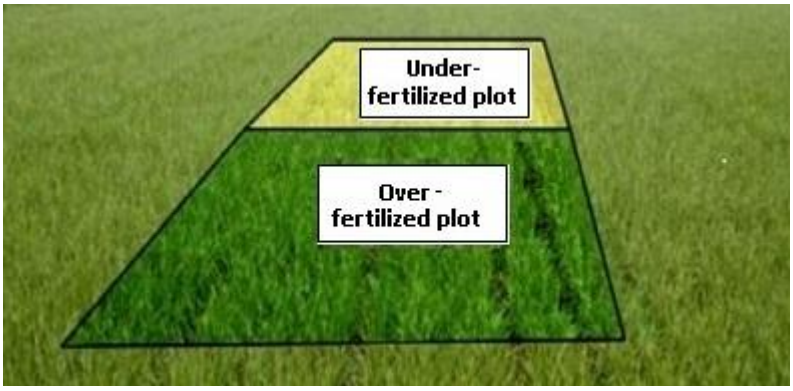


Figure 1-4. An under and over-fertilised plot in the middle of a parcel (Source: modified from Kulmala, 2012).

On this respect Khosla et al. (2005) stated that nitrogen uptake across production fields, was shown to exhibit significant spatial variability. The pattern and scale of the spatial variability was found to be stable over time. As anticipated, site-specific management zones exhibited less N uptake and spatial variability within individual zones than on a whole field basis. Between management zones, N uptake was statistically different. Furthermore, these results are encouraging because development of improved N application algorithms for site-specific management must be consistent over time.

Thrikawala et al. (1999) considered that site-specific management recognises the inherent spatial variability associated with most fields under crop production. According to Koch and Khosla (2003) this new and developing discipline incorporates advanced technologies, such as Variable Rate (VR) fertilisation to enhance the efficiency of farm fertilisers in a profitable and environmentally sensible manner. VR application is one of the most widely used precision technologies. However, the effectiveness of VR fertilisation has so far been limited because nutrient levels in the soil generally show a high within-field variation at scales which change across fields. This spatial variability arises through complex interactions between soil-forming and management factors. Therefore, assessment of the spatial variability of soil fertility parameters is essential for efficient variable-rate application of fertiliser. Conventional fertiliser recommendations are mostly based on soil sampling and laboratory analyses of soil fertility parameters. This procedure is expensive, time consuming and based only on few scattered readings, ignoring within field variability of soil fertility parameters. Consequently, there is a need for fast, cost effective methods for describing within field soil variability (Mouazen, 2011).

1.6. Precision agriculture

Precision agriculture, also known as site-specific management, addresses spatial variability within a field and how to best manage that variability to maximize production and profitability while minimising risk. Site-specific management may be applied to such decisions as variety selection, weed and pest management, nutrient management, and irrigation (Cropsci, 2012).

Sinfield et al. (2010) state that the ever-increasing prices of fertiliser; predominantly nitrate, phosphate and potassium, and the growing ecological concern over chemical run-off into sources of drinking water have brought the concepts of precision agriculture and site-specific management to the forefront of present day technological development within agriculture and ecology. Precision agriculture was initiated in the mid 1980's, using newly available technologies, to improve the application of fertilisers by varying rates and blends as needed within fields in order to address inputs correctly to soil features (Robert, 2002). As underlined by Haneklaus and Schnug (1998), it can provide farmers with a better understanding of the variability in soil fertility within a field and the need for adjustment of the fertiliser application rates to it. This concept has been adapted to a variety of practices, crops, and countries.

Adrian et al. (2005) considered that demographic and economic research perspectives have given a great deal of attention in recent years to the adoption of precision agriculture and the potential for creating efficiencies which are possible with precision agriculture tools. However, Robert (2002) pointed out several agronomical challenges like the lack of basic information, inadequate sampling and scouting procedures, absence of site-specific fertiliser recommendations, misuse of information, and lack of qualified agronomic services. In addition, Robinson et al. (2009) reported that the effects of seasonal and spatial variability in yield maps for precision agriculture are poorly understood and thus may lead to low predictability of future crop yield. Võsa et al. (2009) also noted the large variation in crops growth conditions within a field.

Its implementation, as estimated by Camilli et al. (2007), demands intensive data manipulation no matter how this implementation is effected. Data for some variables can be obtained by means of remote sensing techniques such as satellite images and aerial photos. For others, however, direct field measurements are essential, and this has motivated a great deal of research on sensors, devices, and equipment. Although in some cases field data-acquisition is made with stationary equipment, such as weather stations and with portable devices that are operated manually (such as those for different types of crop scouting), most of the time these sensors and related equipment are embedded in agricultural machines. For example, Ben-Dor et al. (2002) investigated the potential of airborne hyperspectral imagery for mapping soil properties, such as organic matter, field moisture, saturated moisture and salinity. According to these authors the prediction performances obtained for these soil properties were favourable with (R^2>0.65). Moran et al. (1997) reviewed the potential of image-based remote sensing to provide spatially and temporally distributed information for precision crop management (PCM) and found that both aircraft- and satellite-based remote sensing could provide valuable information for PCM applications. López-Granados et al. (2005) compared various prediction methods for mapping soil properties such as organic matter, pH, phosphorus and potassium by incorporating secondary spatial information into the mapping obtained from an intensive grid soil sampling and digital data acquired from an aerial colour photograph of bare soil in combination with geospatial technologies. Adekayode et al. (2009) generated digital maps of the farmland which allow to periodically reviewing the fertility status.

1.7. Vis/NIR spectroscopy: a promising technique for on-site and on-line soil analysis

As Vis/NIR spectroscopy offers a wide range of in-line and at-line transmission and diffuse-reflection probes designed for the measurement of liquids and solids (Siesler, 2002) it is already widely used in industry due to its simplicity, rapidness, and the need for little or no sample preparation (Chang et al., 2001). Therefore, Vis/NIR spectroscopy is being increasingly considered as a possible alternative to complement or replace conventional chemical laboratory methods and to be used for on-site and on-line soil analyses. On-site monitoring of soil nutrient concentration would offer the opportunity for higher density of measurements at relatively lower costs. This would allow for an efficient mapping of nutrient variability to facilitate variable-rate nutrient application. As a result, an implementation of nutrient management programs using sensor technology potentially promotes environmental stewardship while maintaining crop productivity and profitability. Hence it allows significant advances in the area of food production with the application of agricultural inputs, such as fertilisers, seeds, pesticides, among other, in variable rate inside a field, according to the requirements or productive potential of several homogeneous sectors.

Numerous researchers and manufacturers have attempted to develop sensors for measuring soil properties on-the-go. These sensors have been based on electrical and electromagnetic, optical and radiometric, mechanical, acoustic, pneumatic, and electrochemical measurement concepts (Adamchuk et al., 2004; Erickson, 2004; Mouazen et al., 2005; Christy, 2008; Mouazen et al., 2007; Adamchuk and Viscarra Rossel, 2010; Schirrmann and Domsch, 2011; Schirrmann et al., 2011 and Bah et al., 2012).

So, in the context of precision agriculture Vis/NIR spectroscopy might be an alternative for the conventional analysis methods employed in the Villa Clara province for determining soil fertility parameters (OM, P and K_2O) using a single spectrum per sample. In order to achieve this purpose, first it is necessary to calibrate and validate measured laboratory Vis/NIR spectra against laboratory determined soil fertility parameters. As Vis/NIR radiation is mostly absorbed by O-H, C-H and N-H bonds, the organic matter content in soil is expected to have a high impact on the soil spectra. However, some other important soil fertility parameters, such as K_2O and P are not spectrally active in the Vis/NIR range and thus cannot be directly determined from the Vis/NIR spectra of soil samples (Ben-Dor et al., 1997; Malley et al., 2002; Saeys and Ramon, 2004). So, their prediction should be based on their correlation with or

impact on the spectral response of other molecules. If the major soil fertility parameters could be estimated from the soil spectra, Vis/NIR spectroscopy could be a useful tool for on-site analyses of soil samples for soil fertility mapping as input to variable rate fertilisation. If successful this research could promote precision fertilisation in the Villa Clara province. Therefore, a tool which allows the assessment of the variation in soil fertility within a field would increase the efficiency of the applied fertiliser. By replacing the wet chemical analyses by Vis/NIR spectroscopic scanning in the laboratory, it could be possible to speed up the analyses and thus increase the number of samples that could be be analysed. Consequently, it would contribute to reducing the production costs and the negative impacts on the environment related to nutrient losses from the field. Moreover, if it would be possible to bring the Vis/NIR instrument to the site of the field (on-site analyses) it would become feasible to analyse more samples per field. Moreover, by determining the wavelength intervals in the Vis/NIR range that contribute most to the prediction models for the different soil fertility parameters, it might even be possible to develop a lower cost, portable instrument for on-site use.

1.8. Hypothesis, objectives and outlines

According to the previous statements, this dissertation is based on the hypothesis that Vis/NIR spectroscopy could provide a useful tool for on-site measurement of soil fertility parameters for soil fertility mapping as input to variable rate fertilisation.

As it has been shown that the strong Vis/NIR absorption by water complicates the quantification of other molecules, it was decided to work on dried soil samples. Therefore, the main objective of this PhD study was to evaluate the potential of Vis/NIR spectroscopy for the accurate prediction of soil fertility parameters on dried samples from agricultural fields in the Villa Clara province. To accomplish this main objective, three sub-objectives were defined:

1. To test the accuracy of Vis/NIR spectroscopy for measuring the total P level in soil when it is not correlated to the organic matter content.
2. To assess the potential of Vis/NIR spectroscopy for measuring the average fertility values for soil samples at landscape level in the Villa Clara province.
3. To validate the Vis/NIR spectroscopy technique as a rapid alternative for predicting the variation in soil fertility parameters within a field in the Villa Clara province.

Chapter 2

State of the art on Vis/NIR spectroscopy for soil analysis

2.1. Introduction

Historically, our understanding of the soil and assessment of its quality and function has been obtained through routine soil chemical and physical laboratory analyses as underlined by Viscarra Rossel et al. (2006). The variability in space and time for point data can give valuable insight into the dynamic nature of soil fertility parameters within a field's boundary (Cox et al., 2003). According to Christy (2008), characterizing the spatial variability of soil attributes using traditional soil sampling and laboratory analyses would be too expensive to be feasible. However, precision agriculture is based on the application of agronomic principles using advanced technologies to manage spatial and temporal variability related to the agricultural production process, for instance to characterize the variability in the soil properties within one or more fields in a relatively small geographical area (Adamchuk et al., 2004 and Volkan et al., 2010).

Precision fertilisation requires a finely meshed knowledge of how these soil properties vary in the individual fields to obtain the economic and environmental advantages of the technology (Sorensen and Dalsgaard, 2005), but the high cost of soil sampling and chemical analyses by conventional laboratory methods has so far restricted the full implementation of this technique at field level. Therefore, Meyer et al. (2005) and Lv and Yang (2011) considered the application of Visible and Near Infrared Reflectance (Vis/NIR) spectroscopy as a nondestructive testing method with advantages of repeatability, time-saving and cost-effectiveness. This method has the potential to facilitate a rapid assessment of soil fertility parameters. The operating principle of Vis/NIR spectroscopy is based on sending light in the wavelength range 400 – 2500 nm onto the sample and then measuring the light coming from the sample at different wavelengths. This covers the visible (380 – 780 nm) and near infrared (780 – 2500 nm) regions (Rehbein and Oehlenschlager, 2009). The absorption of light in the 700 to 2500 nm region is primarily

by O-H, C-H and N-H bonds, so it is especially useful for measuring organic substances.According to Heise and Winzen (2002) there is a tendency to perform fewer experiments, while recording more and more data during each of them. Dessipri et al. (2003) pointed out that its use to monitor various processes has been developed systematically during the last decades. Based on the previous considerations this chapter aims to evaluate the applicability of Vis/NIR for prediction of soil fertility parameters in the context of precision agriculture based on the studies reported in literature. First, the general concepts of precision agriculture, Vis/NIR measurements in laboratory conditions and the building of calibration models by means of multivariate statistical methods, linking the results of the soil chemical and spectral analyses, are presented. In the main part of the chapter, the accuracies of calibration models obtained in laboratory conditions and with portable Vis/NIR spectrophotometers as reported by different researchers are reviewed. Finally, conclusions are drawn with respect to the applicability of Vis/NIR spectroscopy for soil analyses in the context of precision agriculture.

2.2. Visible and Near Infrared Reflectance Spectroscopy (Vis/NIR)

Most of the work on spectral analyses for quantifying soil properties has been focused on Vis/NIR spectroscopy, a simple and non-destructive analytical method that can be used to predict several properties simultaneously (Chang et al., 2001; McCarty et al., 2002; Viscarra Rossel et al., 2006; Nduwamungu et al., 2009; Volkan et al., 2010; Stenberg et al., 2010). Vis/NIR spectroscopy has been used since the early fifties for chemical analyses, especially in the field of agriculture. However, the real breakthrough as a quality- and process-control tool has occurred only within the last two decades, since the introduction of efficient multivariate calibration techniques and the development of light-fibre coupled probes.

Siesler (2002) stated that the information contained in NIR spectra is less specific than that in visible, ultraviolet (UV) or infrared spectroscopy, because the absorption in this range is dominated by overtones and combinations of the fundamental molecular vibrations. On the other hand, thanks to the lower molecular absorptivity compared to the other wavelength ranges, larger optical path lengths are feasible such that minimal sample preparation is needed. This makes it possible to acquire NIR spectra under difficult conditions, such as in the field or on the factory line. Consequently, despite the

lack of comparable specific spectral information, NIR spectroscopy is quickly overtaking Raman spectroscopy as a process-monitoring technique.

Due to the lack of highly specific spectral information, He et al. (2005) described this technique as basically an empirical method, whose complete procedure consists of two stages: (set I) calibration stage where the prediction equation is developed and (set II) validation stage where the previous stage is validated. This technique shows many possibilities in the field of testing chemical and physical properties of materials.

2.3. Vis/NIR spectrophotometers

In terms of instrumentation Vis/NIR spectroscopy is not distinctively different from either UV or infrared spectroscopy, although it may sound unique (Kawata, 2002). Traditionally, dispersive instruments, available since the 1940s, were used to obtain infrared spectra. With the introduction of light-fiber optics in the mid-1980s and the monochromator-detector developments in early-1990s, Vis/NIR spectroscopy became a more powerful tool for scientific research. The wavelength range is one of the main components affecting the measurement accuracy with Vis/NIR spectroscopy (Mouazen et al., 2006). Instrumentation for Vis/NIR spectroscopy includes a source, a detector, and a dispersive element to allow the intensity at different wavelengths to be recorded. Fourier-transform infrared spectrometers using an interferometer are now frequently used and have improved the acquisition of infrared spectra dramatically (Stuart, 2004), especially for wavelengths above ~1000 nm. Other Vis/NIR instruments are single monochromators and spectrographs, which according to Forment and Hanselaer (2010) are not able to provide satisfying accuracy due to poor stray light rejection. Therefore, double monochromators are often used. The main drawbacks of a double monochromator scanning system are the need for motion synchronization and the time-consuming scanning approach, which make it inconvenient for on-site and field measurements. Mouazen et al. (2006) evaluated the performance of two commercially-available spectrophotometers with different wavelength ranges [diode array with a short wavelength range (SWR) of 300-1700 nm and a combination of a diode array and scanning monochromator with a full wavelength range (FWR) of 350-2500 nm] for measurement of selected soil attributes. They concluded that under wet field conditions pH, available phosphorous (P), cation exchange capacity (CEC), potassium (K) and

calcium (Ca) were more accurately predicted with the SWR, whereas total nitrogen (N), total carbon (C), sodium (Na) and magnesium (Mg) were better predicted with the FWR.

2.4. Sample presentation

Depending on the sample, the spectrum can be measured in either reflection or transmission. Vendors of Vis/NIR spectroscopy equipment provide a wide range of in-line and at-line transmission and diffuse-reflection probes designed for the measurement of liquids and solids, which make it a very practical technique for rapid and non-destructive analyses. As water is a much weaker absorber in the NIR than in the MIR and UV, NIR spectroscopy can also be used for measuring aqueous solutions in cuvettes of a few millimetres. Sample information in the near infrared region is usually collected as an absorption spectrum through transmission measurements or diffuse-reflectance measurements with a Vis/NIR spectrophotometer. Reflectance methods can be divided into two categories: internal reflectance measurements in contact with the sample and external reflectance measurements which involve a light beam reflected directly from the sample surface (Stuart, 2004).

In Vis/NIR measurements, diffuse reflection or scattering of light at the sample is very commonly detected. In both cases, the surface of the sample is illuminated and the light is scattered at or in the sample into a wide range of angles, so that introduction of all the light into the spectrometer is difficult (Kawata, 2002). For instance, Volkan et al. (2010) found that soil reflectance decreases with increasing soil organic matter and Viscarra Rossel et al. (2008) showed their results with indirect measurements of soil organic carbon (OC) and iron (Fe) contents using soil colour as the proxy. Black and dark colours absorb more light than light colours or white. As such, dark soils tend to be warmer than light coloured soils when the sun shines or when the atmosphere is warm and the soil is able to absorb energy from it (Akinrinde, 2006).

2.5. Interferences

NIR reflectance is quite sensitive to external environmental conditions, such as temperature and soil moisture (Minasny et al., 2011). When samples that have different temperatures from that of the calibration sample set are predicted, a bias necessarily occurs (Kawano, 2002). According to Wülfert (2004) in most cases temperature measurements are readily available at the same time that spectra are measured. Since the temperature is therefore a known quantity it can be appendedd to the spectra and used as

independent variable in order to improve prediction. This author found that the collinearity between temperatures affected regions of the spectrum modelling and it appears that it is not possible to correct for temperature influences with linear techniques because of the non-linear character of the effects like shape changes caused by an influence on the long range intermolecular and intramolecular forces such as H-bonding.

In this regard, depending on the topography of an area and the soil characteristics, soil moisture can be quite variable over the land surface. Thus, a method for determination of soil moisture without the necessity for exhaustive manual measurements would be beneficial in characterizing soil moisture within a given region or field (Kaleita et al., 2005). In quantitative analyses, as underlined by Kawano (2002), a variation in moisture content may be accepted if it is within the moisture range of the calibration samples, but it is important to keep the moisture content of samples within this range. Otherwise, a bias may occur. In order to avoid these problems, it is better to perform a temperature and moisture control of the samples. For instance, Wülfert (2004) reported that explicit inclusion of the temperature into the model is often expected to improve the accuracy. Norris and Davies (2012) found that using the fourth derivative with a carefully chosen gap can enhance NIR spectra and appears particularly valuable for understanding the effects of sample temperature on Vis/NIR spectra. Minasny et al. (2011) utilized an external parameter orthogonalisation (EPO) algorithm to remove the effect of soil moisture from NIR spectra for the calibration of soil organic carbon content. Lobell and Asner (2002) explained that a volumetric moisture range of 0 to 50% spans the most common conditions found in arid and semi-arid ecosystems and in other areas such as dormant agricultural fields. Thus, these authors sought to quantify changes in soil reflectance as a soil proceeded from wet to dry state and to determine the dependence of these changes on soil type and wavelength. They found that the reflectance decreased with increasing moisture content for all soils and these observed changes in soil reflectance revealed a nonlinear dependence on moisture that was well described by an exponential model and was similar for different soil types when moisture was expressed as degree of saturation. The simple moisture reflectance model obtained by these authors that required only dry soil reflectance as input, demonstrates the potential for monitoring moisture conditions in exposed soils from airborne and spaceborne vantage points and this model might also be useful for constraining soil reflectance in spectral unmixing algorithms and radiative transfer models to derive canopy characteristics.

Chang et al. (2001) evaluated the ability of Vis/NIR spectroscopy to predict various soil properties and investigated why Vis/NIR spectroscopy can be used to predict some soil properties that theoretically do not respond to light in the near infrared region. These authors found that all soils tested had similar Vis/NIR reflectance spectra characterized by a high optical density [log (1/R)] in the visible light region (400~800 nm) and three major absorption peaks (around 1400, 1900, and 2200 nm) in the near infrared light region. They concluded that the accurate predictions obtained on these soil properties (cation-exchange capacity and 1.5 MPa water), without a primary response in the NIR region, could be explained by the correlation with clay and organic matter.

2.6. Chemometric analyses

An NIR spectrum is commonly obtained by passing infrared radiation through a sample and determining what fraction of the incident radiation is absorbed at a particular waveband (Stuart, 2004). NIR spectra are dominated by the weak overtones and combinations of C-H, N-H, and O-H bonds for which the fundamental vibrational bands are situated in the mid infrared region. As these overtone and combination bands result in broad and overlapping absorption peaks, this makes NIR spectra difficult to interpret (Chang et al., 2001; McCarty et al., 2002). These divergent spectral characteristics may be expected to have substantial influence on the ability to obtain quantitative information from spectral data. Therefore, advanced multivariate statistical methods are needed to capture the relation between spectral variation and the variation in the components/properties of interest.

Chemometrics is the discipline that uses mathematical and statistical methods for the selection of the optimal experimental procedures and the appropriate treatments of the data in chemical analyses and includes three main groups of techniques compiled as (i) pre-treatment (pre-processing), where spectroscopic reflectance is pre-processed (Nduwamungu et al., 2009), (ii) classification methods, to group samples according to their spectrum and (iii) regression methods, to relate the spectrum with some quantifiable properties of the samples.

Spectral pre-processing techniques are used to remove any irrelevant information which cannot be handled properly by the modelling techniques (Schwartz et al., 2011). The common sources of this undesirable information are baseline shift, light scattering, path length differences and instrumental drift. Therefore, spectral pre-treatments or

pre-processing techniques, which are mathematical functions, have to be used to handle or remove such interferences from the raw spectral signals. The main groups of pre-processing include spectral transformations for linearization [log (1/R), log (1/T), Kubelka-Munk], scatter correction techniques [baseline correction, detrend, derivatives, smoothing, multiplicative scatter correction (MSC), extended multiplicative signal correction (EMSC)], interference removal [orthogonal signal correction (OSC), external projection orthogonalisation (EPO)], weighting of variables [standard normal variate (SNV), generalized least squares weighting (GLSW)] and variable centering (Mean Centre). The most commonly used pre-processing techniques for Vis/NIR spectroscopy on soil are log (1/R), derivatives, smoothing, MSC, normalisation, SNV and Mean Centre.

The most used linear transforms include log (1/R), which is commonly called absorbance, where R is relative reflectance. If a matrix absorbs at the same wavelengths as the analyte, log (1/R) will prove a better choice for relating reflectance to concentration. This transformation is basically well suited for diffuse reflectance NIR spectra of mixtures with absorbing matrices. Plots of R versus concentration are typically less linear than plots of log (1/R) versus concentration (Ulmschneider and Roggo, 2008).

Derivatives are usually calculated according to the Savitzky-Golay algorithm and typically only the 1st and 2nd derivative are used. The 1st derivative is often used to remove baseline shifts and superposed peaks and 2nd derivative can correct locally for both additive and multiplicative effects (Naes et al., 2002). In order to avoid noisy derivative spectra some smoothing has to be incorporated in the derivative calculation. Therefore, by applying smoothing it is possible to reduce the random noise in the instrumental signal and keep useful variation (Schwartz et al., 2011). However, the multivariate calibration techniques also perform some smoothing action and thus the added value of spectral smoothing in terms of prediction performance is usually very limited.

MSC is a scatter correction method where the light scattering is estimated for each sample spectrum relative to a reference spectrum, which is typically the mean spectrum. Each spectrum is then corrected such that all samples appear to have the same scatter level as the reference spectrum.

OSC is a spectral treatment that allows the portion of spectral information orthogonal to changes in analyte concentrations to be removed. OSC effectively removes information

not correlated to the target parameter, which substantially decreases the number of PLS (Partial Least Squares) components required to construct calibration models. This, however, does not increase predictive ability because not all variability not correlated to the concentration of the target analyte is orthogonal to it (Bertran et al., 2001). EPO deals with the case where the external parameter can not be measured online. Spectra are pre-processed by projection onto the orthogonal to the space in which alterations induced by the external parameter variations occur (Roger et al., 2003).

Normalisation is typically used to obtain all data into approximately the same scale or to obtain a more even distribution of the variances and the average values (Barnes, et al., 1989). Normalisation can be classified into the maximum, mean and range normalisation. The maximum normalisation is an alternative process which divides each reflectance by the absolute value of maximum reflectance available in each row of the spectrum matrix. The maximum value becomes +1 and the minimum value becomes -1. The mean normalisation consists of dividing each row of a spectrum matrix by its average, thus neutralizing the influence of the hidden factor. Finally, in range normalisation, each row is divided by its range; the range is the maximum value minus minimum value (Cogdill, 2008).

SNV removes the slope variation from spectra caused by scatter and variation of particle size (Barnes et al., 1989). The transformation is applied to each spectrum individually by subtracting the spectrum mean and scaling with the spectrum standard deviation (Jørgensen, 2000). Mean Centre is one of the most common pre-processing methods, which calculates the mean of each column and subtracts this from the column. Another way of interpreting Mean Centre data is that, after Mean Centre, each row of the mean centred data includes only how that row differs from the average sample in the original data matrix (Eigenvector, 2012).

The **regression methods** include the Principal Component Regression (PCR) which uses Principal Component Analyses (PCA) to project the original variables onto a smaller number of orthogonal linear combinations (principal components) prior to estimating the regression coefficients. This is done to avoid the multi-collinearity problems which arise from having a large number of highly correlated wavelength variables. Partial Least Squares (PLS) regression applies a similar projection onto orthogonal linear combinations, but searches the linear combinations (latent variables) which explain

maximal covariance between the spectra and the component of interest, rather than searching those that explain maximal variance in the spectra, as PCA does. So, unlike PCR, PLS balances the two objectives of explaining response and predictor variation (Viscarra Rossel et al., 2006). PLSR is the most commonly used multivariate statistical method in calibrating soil reflectance to individual soil parameters (Volkan et al., 2010). In this sense Viscarra Rossel (2007) reported that the robustness of PLS models and their predictions may be further improved by combining the implementation of PLS with bootstrap aggregation or "bagging", which aims to reduce the variance of predictions by aggregating a number of models obtained in the course of re-sampling.

When the relation between the spectral variation and the component of interest is non linear, non linear regression techniques such as Locally Weighted Regression (LWR), Support Vector Machine (SVM) or Artificial Neural Networks (ANN) can be used (Twarakavi et al., 2009; Igne et al., 2010; Scott et al., 2011). Also reformulations to the standard SVM, like Least Squares Support Vector Machines (LS-SVM), can be used (Suykens et al., 2002; Zhang et al., 2008; Oliveira et al., 2010; Kovačević et al., 2010; and Prasad et al., 2011). For instance, Shao and He (2011) applied Least-squares support vector machines (LS-SVM) to construct calibration models for soil properties such as available N, P and K and obtained better results with LS-SVM than with standard PLSR.

Donigian (2002) considered model calibration and validation like necessary and critical steps in any model application. Calibration is an iterative procedure of parameter evaluation and refinement, as a result of comparing simulated and observed values of interest. Model validation is in reality an extension of the calibration process. Its purpose is to assure that the calibrated model properly assesses all the variables and conditions which can affect model results, and demonstrate the ability to predict field observations separate from the calibration effort.

Model performance and calibration/validation are evaluated through qualitative and quantitative measures, involving both graphical comparisons and statistical tests. Model performance is often based primarily on visual and graphical presentations. Several **statistics** can be used to quantify the predictive ability of the calibration models. The coefficient of determination R^2 measures the proportion of the total variation in the Y-variable accounted for by the model. The RMSE (root-mean-square error) is the absolute

standard error on the predictions which can be calculated for the calibration samples (RMSEC) or the test set prediction (RMSEP).

2.7. Vis/NIR spectroscopy for soil analyses in a precision agriculture context

Several researchers have investigated the potential of Vis/NIR spectroscopy in a precision agriculture context for quantifying important soil chemical properties related to fertility, such as the content of organic matter (OM), organic carbon (OC), total carbon (TC), phosphorus (P) and potassium (K), and the pH. An overview of the results obtained in different published studies on soil properties prediction with Vis/NIR spectroscopy using dried samples is given in Table 2-1. These results are variable and the differences in obtained accuracy can probably be attributed to differences in sampling strategy (e.g., geographic range of the studies), sample preparation (e.g., sieving versus grinding, air drying versus over drying), spectral acquisition procedure, spectral pre-processing, calibration procedure and reference method (Ndwamungu et al., 2009). In general the best results were obtained for OM, OC and TC, while the results for P, K and pH varied from very poor to reasonably good.

For instance, Lee et al. (2003) conducted a study to develop fundamental relationships from representative soil orders (the highest level of soil classification) and their spectral characteristics. They reported that the mean pH, soil organic matter, and P values among these soil orders (Alfisol, Entisol, and Ultisol) were significantly different at the 0.05 level, but the mean K was not significantly different (α= 0.05) among the soil orders. They also found that the loadings for the different calibration models were different. This indicates that the Vis/NIR calibrations were using different information for the different P forms or associations of P with soil organic matter and other components of the soil matrix as assessed by the different reference methods.

He et al. (2005) evaluated the potential of Vis/NIR spectroscopy on mixed loamy soil samples. For nitrogen (N) and organic matter (OM) they obtained higher correlation coefficients (r) between measured and predicted values of 0.92 and 0.93, as well as low standard errors of prediction (SEP) of 3.29% and 0.07%, respectively. This indicates that Vis/NIR has potential to accurately predict these constituents in this soil. However, the prediction results they obtained for phosphorus (P) and potassium (K) were less encouraging with correlation coefficients of 0.47 and 0.69, and SEP values of 33.38% and 25.02%, respectively.

Volkan et al. (2010) attributed the poor prediction results for pH and K to the narrow chemical range, high skewness of these variables in the data sets and a lack of information on pH in the Vis/NIR spectra. Also, the predictions for exchangeable K obtained by Islam et al. (2003) were poor. However, these authors were able to successfully predict pH and OC using principal component regression (PCR) on randomly selected calibration and validation sets. As well, Thomsen et al. (2009) and Ladoni et al. (2010) reported that Vis/NIR spectroscopy was able to predict the soil OC. Several other researchers (Dalal and Henry, 1986; Stevens et al., 2008; Vasques et al., 2009; Viscarra Rossel and Behrens, 2010) also reported on the suitability of Vis/NIR spectroscopy to quantify the organic carbon content in soil. Bogrekci and Lee (2005a) reported R^2 values of 0.78; 0.87, and 0.92 for the prediction of the P content in soils from the same acquired NIR spectra, but with different reference methods for measuring the P content. This indicates that Vis/NIR spectroscopy and the reference methods are not necessarily measuring the same P fraction. Morón and Cozzolino (2007) evaluated the effect of two reference methods, namely Bray and Resins, on the Vis/NIR calibrations to predict phosphorus in soil samples. They also reported an effect of the reference method for determining the P content in soil on the prediction accuracy of Vis/NIR spectroscopy.

Table 2-1

Vis/NIR validation performance of soil chemical properties with different reference methods.

Soil property	Reference method	Soil sample condition	Chemometric analyses	Spectral range	Validation performance		Soils location	Sources
					R^2 [a]	RMSEP [b] (%)		
OM	Walkley-Black	Oven dried	PLS	400 – 2498	0.49	0.23	Florida, USA	Lee et al. (2003)
OM	Walkley-Black	Air-dried	PLS	350 – 2500	0.93	0.13	Zheijang, China	He et al. (2005)
OM	Walkley-Black	Air-dried	PLS	350 – 2500	0.93	0.12	Zheijang, China	He et al. (2007)
OM	Walkley-Black	Air-dried	PLS	350 – 2500	0.75	2.90	Turkey	Volkan et al. (2010)
OC	Wet oxidation	Air-dried	PLS	400 – 2500	0.86 – 0.93	-	Canada	Ndwamungu et al. (2009)
OC	-	Air-dried	PLS	320 – 1000	0.83	-	Brittany, France	Aïchi et al. (2009)
OC	Walkley-Black	Air-dried	PCA	700 – 2500	0.76	-	-	Islam et al., 2003
TC	Dry combustion	Air-dried	PLS	400 – 2500	0.79 – 0.96		Canada	Ndwamungu et al. (2009)
P	Olsen	Air-dried	PLS	350 – 2500	0.47	33.38	Zheijang, China	He et al. (2007)
P	Mehlich-I-extractable	Oven dried	PLS	400 – 2498	0.72	18.60	Florida, USA	Lee et al. (2003)
K	Richards (1954)	Air-dried	PLS	350 – 2500	0.67	49.99	Zheijang, China	He et al. (2007)
K	Mehlich-I-extractable	Oven dried	PLS	400 – 2498	0.24	-	Florida, USA	Lee et al. (2003)
K	0.01 M silver-thiourea	Air-dried	PCA	700 – 2500	0.00	-	Australia	Islam et al. (2003)
K	NH_4OAc extraction	Air-dried	PLS	350 – 2500	0.38	0.20	Turkey	Volkan et al. (2010)
pH	1:2 soil: H_2O	Air-dried	PLS	400 – 2500	0.49 – 0.74	-	Canada	Ndwamungu et al. (2009)
pH	1:5 soil: H_2O	Air-dried	PCA	700 – 2500	0.71	-	-	Islam et al., 2003
pH	1:2 soil: H_2O	Oven dried	PLS	400 – 2498	0.78	-	Florida, USA	Lee et al. (2003)

[a] R^2 - coefficient of determination; [b] RMSEP- root mean square error of prediction;
OM (Organic matter); OC (Organic carbon); TC (Total carbon); P (Phosphorus); K (Potassium)

Table 2-2 summarises the results obtained in fresh soil by using an on-the-go soil sensor for on-line spectral measurements for OM (Christy, 2008), Olsen P (Maleki et al., 2006) and other soil properties (Mouazen et al., 2007). Some of these results for OM, organic carbon (TC), total carbon (TC), P and K obtained on-site (fresh soils), showed similarities in R^2 prediction with those achieved in laboratory conditions (air-dried soils, Table 2-2). Therefore, the major benefit of on-the-go sensing would be the ability to quantify the heterogeneity (non-uniformity) of the soil within a field and to adjust other data collection and field management strategies accordingly (Adamchuk et al., 2004).

As new on-the-go soil sensors are developed, different real-time and map-based variable rate soil treatments may become economically feasible. The relatively high costs of soil sampling and laboratory analysis suggest a need for on-the-go soil sensors that could detect critical soil properties in every field location while moving these sensors across the field (Adamchuk et al., 2004). The sensors may be invasive, or not, and may or may not be mounted on vehicles for on-the-go operation (Adamchuk and Viscarra Rossel, 2010).

At a lower cost than using the traditional methods, Adamchuk (2011) considered the possibility for obtaining high-density soil measurements, which are produced using four main types of sensors. Unfortunately, these data may not be adequate for an appropriate decision support system as the measurements may react to a change of more than one agronomic soil property. Therefore, future research should focus on: 1) sensor fusion to bring together sensing components that may have different degrees of response to different soil phenomena, 2) localized sensor calibration methods to define relationships between sensor outputs and laboratory soil test results for specific environments, 3) data integration to employ benefits of remote sensing, proximal crop canopy sensing and yield mapping to better understand manageable soil processes, and 4) turn-key applications to make sensor technology accessible, affordable and useful for agricultural production, land remediation and other situations where soil heterogeneity is an influential factor.

Table 2-2

On-line Vis/NIR validation performance of soil fertility parameters with different references methods.

Soil property	Reference method	Soil sample condition	Calibration method	Spectral range	Validation performances		Soils location	Sources
					R^2 [a]	RMSEP [b] (%)		
OM	Loss on ignition (LOI)	Fresh	PCR	950 - 1650	0.80	0.40	Kansas, USA	Christy (2008)
OM	Loss on ignition (LOI)	Fresh	PCR	950 - 1650	0.75	0.44	Kansas, USA	Christy (2008)
OM	Loss on ignition (LOI)	Fresh	PCR	950 - 1650	0.67	0.52	Kansas, USA	Christy (2008)
OC	Walkley-Black	Fresh	PLS	307 – 1711	0.74	0.48	Belgium, France	Mouazen et al. (2007)
TC	Dry combustion	Fresh	PLS	307 – 1711	0.73	0.27	Belgium, France	Mouazen et al. (2007)
P	Olsen	Fresh	PLS	307 – 1711	0.69	1.35	Belgium, France	Mouazen et al. (2007)
P	Ammonium lactate	Fresh	PLS	307 – 1711	0.73	11.52	Belgium, France	Mouazen et al. (2007)
P	Olsen	Fresh	PLS	401 – 1663	0.68	-	Flanders, Belgium	Maleki et al. (2006)
P	Olsen	Fresh	PLS	401 – 1663	0.67	-	Flanders, Belgium	Maleki et al. (2006)
P	Olsen	Fresh	PLS	401 – 1663	0.63	-	Flanders, Belgium	Maleki et al. (2006)
P	Olsen	Fresh	PLS	401 – 1663	0.63	-	Flanders, Belgium	Maleki et al. (2006)
K	Mehlich I extraction/ICP	Fresh	PCR	950 - 1650	0.80	0.40	Kansas; USA	Christy (2008)

[a] R^2 - coefficient of determination; [b] RMSEP- root mean square error of prediction;
OM (Organic matter); OC (Organic carbon); TC (Total carbon); P (Phosphorus); K (Potassium)

2.8. Conclusions

The potential of Vis/NIR spectroscopy as a rapid and versatile technique for soil fertility assessment in the context of precision agriculture has been investigated by reviewing the state of the art. According to the references consulted in this chapter Vis/NIR spectroscopy can successfully predict soil fertility parameters such as organic matter, organic carbon and total carbon with high coefficients of determination (R^2) ranging from 0.75 and 0.93 for air-dried soil. For on-line measurements on fresh soil, the prediction performance for organic matter was considerably worse with R^2 values ranging from 0.67 to 0.80. The results for other properties such as potassium, phosphorus and pH the results were lower than for OM and varied considerably between different studies. Therefore, it can be concluded from this review that Vis/NIR spectroscopy definitely has potential for predicting the organic matter content in air-dried soil samples better than in on-line measurements performed on fresh soil. This technique may also be used for predicting the P and K content in air-dried soil samples with satisfactory results. As such, Vis/NIR spectroscopy could be a useful nondestructive technique for measuring the fertility of Cuban agricultural soils and could even allow to assess the local soil fertility as an input for site specific fertilisation and crop production management.

Chapter 3

Spatial variability and correlation of soil fertility parameters

3.1. Introduction

Conventional agriculture treats an entire field uniformly with respect to the application of fertilisers. However, soil is spatially heterogeneous, with most soil fertility parameters varying significantly within just a meter. Thus, soil spatial variability is one of several factors that cause within-field variation in crop yield (Corwin, 2005; Corwin and Lesch, 2008; Patil et al., 2011). Characteristics and variability of different soil fertility parameters have been reported, analysed and detailed in several scientific sources (Pleysier, 1995; Brouder et al., 2001; Braud et al., 2003; van Schöll and Nieuwenhuis, 2004). Knowledge of the spatial variability of these soil fertility parameters and analyses of the spatial correlation between them, are important for sugarcane production. Recognising the significance of quantifying and managing this variability that occurs in agricultural fields, could lead to different approaches for implementing a site-specific management in sugarcane production in the Villa Clara province. Also, these are the most commonly analysed parameters within the SERFE system.

Organic matter

Organic matter only constitutes a few percent of the soil weight (1 – 6%), but its quality and quantity influences almost all the properties which contribute to soil fertility (Magdoff, 1997). For simplicity, organic matter can be divided into two major categories: stabilized organic matter which is highly decomposed and stable, and the active fraction which is being actively used and transformed by living plants, animals, and microbes. Two other categories of organic compounds are living organisms and fresh organic residue. These may or may not be included in some definitions of soil organic matter (UMN, 2002). The soil organic matter in its different forms has distinct effects on almost all the soil properties. According to Bot and Benites (2005), in terms of improving soil structure, the active and some of the resistant soil organic components, together with micro-organisms (especially fungi), are involved in binding soil particles into macro-aggregates.

According to Six et al. (2004) the aggregate hierarchy concept proposed by Tisdall and Oades (1982) is probably the most significant theoretical advancement in the understanding of aggregate – OM interactions. In the aggregate hierarchy concept it is postulated that the different binding agents act at different hierarchical stages of aggregation. Free primary particles and silt-sized aggregates (<20 μm) are bound together into microaggregates (20–250 μm) by persistent binding agents (e.g., humified organic matter and polyvalent metal cation complexes), oxides and highly disordered aluminosilicates. These stable microaggregates, in turn are bound together into macroaggregates (>250 μm) by temporary (e.g., fungal hyphae and roots) and transient (e.g., microbial- and plant-derived polysaccharides) binding agents. Aggregation is important for good soil structure, aeration, water infiltration and resistance to erosion and crusting. Therefore, balanced soil fertility is crucial for keeping a high quality crop production. Thus, the role of soil organic matter is important for maintaining the productivity of soils.

Phosphorus

The primary source of phosphorus in native soils is apatite, a calcium phosphate that also contains hydroxides or fluorides. Other sources include decaying plant and animal residues, humus, and microorganisms. In acid soils, iron and aluminum in solution and in oxide and hydroxide forms react strongly with added P, rendering it unavailable to plants (Hodges, 2010; Nezat et al., 2008). Phosphorus exists in inorganic and organic form in soil (Olsen and Sommers, 1982). The soil inorganic phosphates result from the combination of the phosphate species with metallic cations which can substitute one or more hydrogen atoms.

Hodges (2010) considered phosphorus as unique among the anions because it has low mobility and availability in soils. It is difficult to manage because it reacts so strongly with both solution and solid phases of the soil. As a result, mobility through the soil is extremely limited in all but organic soils or white bleached sands with extremely low specific surface. While P occurs in a multitude of inorganic and organic form in the soil, the orthophosphates, $H_2PO_4^-$ and HPO_4^{2-}, are the primary forms of phosphorus taken up by plants, with the dominant form determined by the soil pH. In soils with pH values exceeding 7.0 the HPO_4^{2-} form predominates, while in soils with pH between 4.3 and 7.0, the $H_2PO_4^-$ form predominates. Regardless of the form, the concentration of soluble P in

soil solution is very low. Native soil P levels are often low enough to limit crop production (Daniels et al., 2004). The removal of this element from the soil is mainly through extraction by crops and through erosion, because the losses by leaching are minimal, due to its low solubility and its limited movement. Only in sandy and in organic soils the leaching could be significant. On the other hand Fassbender (1975) stated that its high stability in soil results in a low solubility, which sometimes causes deficiencies in phosphorus availability for plants, in spite of the continuous mineralization of soil organic compounds.

Potassium

The availability of soil potassium depends primarily on the types and amounts of soil minerals present. Unavailable potassium found in rocks and primary minerals becomes available only after these minerals begin to weather. Soil clays which can collapse around the potassium ion are capable of "fixing" K so that it is no longer free to exchange with other cations in solution. This pool of potassium is termed slowly available, and can be very important in soils containing vermiculite and some montmorillonite in their clay fractions. Slowly available K is released by acid weathering of the clay or by cycles of wetting and drying. Soil solution and exchangeable K are considered readily available for plant uptake. Soil solution rarely holds more than 8.97 to 11.22 kg ha^{-1} of available K (Hodges, 2010). Satisfactory crop production is obtained only when the soil possesses favourable soil fertility parameters. However, the management based on the average soil fertility (Figure 3-1) is no longer enough to maintain sustainable high yields in the agricultural production.

Figure 3-1. Uniform application of fertilisers within a field
Source: Weisz and El Mehelmy (2012).

Fertiliser can be applied in a number of ways such as broadcasting (the fertiliser pellets are spread evenly over the whole field, and then often ploughed or raked into the soil) and row application (the fertiliser is applied in rows). In this way, several authors such as Olsen *et al*. (1971), Rowse and Stone (1980), and Mandal and Thakur (2010) reported that only 40 to 50% of N fertilisers and 20 to 30% of P and K fertilisers are effectively used by crops. Thus, the remaining % become volatilised, leached to groundwater, or get fixed within the soil. Also, broadcasting of fertilisers, especially P and K, produces fixation problems due to greater soil contact, whereas volatilisation of N results in a decrease in the amount of N available content in the soil. Applying fertiliser in this manner results in lower productivity and profitability due to missing out on additional yield in the parts of the field that are under-fertilised and further reduced profitability where fertiliser is over-applied (Phillips, 2009). For instance, Mandal and Thakur (2010) designed a subsoiler-cum-differential rate fertiliser applicator. The equipment was tested in the field to observe its performance on sugarcane with results showing an increase of 16.2%, 16.4% and 35.4% in yield as compared to conventional ploughing within-furrow fertiliser application.

The productive capacity of the agricultural soils used for sugarcane production in Villa Clara province can be directly related to the improvement of soil fertility parameters such as organic matter, phosphorus and potassium. The knowledge about the spatial variability of these soil fertility parameters can be used to implement a group of suitable agricultural practices for applying fertilisers in these crop systems. This can help to maximize the economic benefits and to minimise the effects of environmental impact. On this respect, according to the SERFE recommendations, the mechanized handling and distribution of fertilisers (NPK), is carried out by using the fertiliser spreader F – 350, equipped with three hoppers and modified to bury the fertiliser to a depth of 10-15 cm (Figure 3-2).

Figure 3-2. Fertiliser spreader F - 350 equipped with three hoppers.

For that reason in this chapter a detailed description of the studied areas, soil sampling, reference analyses and data processing methods is given in section 3.2. The results are discussed in section 3.3. Finally, conclusions are given in section 3.4.

3.2. Materials and methods

3.2.1. Study areas

The studied fields are all located in the Villa Clara province in the central part of Cuba between the coordinates 22°16', 23°09' N and 80°02', 80°25' W. The Villa Clara province is surrounded by the Atlantic Ocean in the north, the Sancti Spiritus and Cienfuegos provinces in the east, south-east and south-west, and the Matanzas province in the west (ONE, 2009). It is sub-divided into 13 municipalities and its capital is Santa Clara city, situated at a junction of the old central road, the national highway and the central railway pass (Figure 3-3).

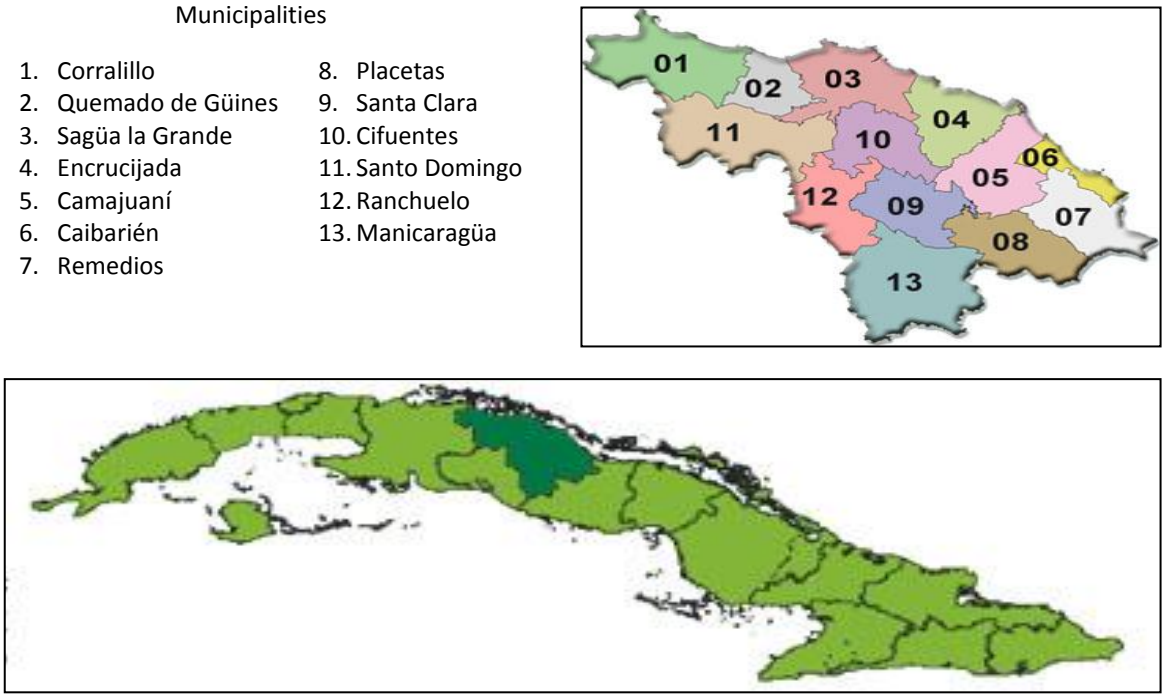


Figure 3-3. Villa Clara province.
(Source: ONEI, 2010)

Its highest point is "Pico tuerto" in the south east of the city, rising 919 m above sea level and merging into the Sierra Escambray. According to ONEI (2010) the province occupies a significant extension of 8 411.81 km^2 which represents 7.76% of the total surface area of the country. At the end of 2010 the population was 800 335 inhabitants.

The research was conducted at 10 Agroindustrial Complexes (CAI) situated in the Villa Clara province. These areas are located in different edaphoclimatic zones (Figure 3-4).

Agroindustrial Complex (CAI)

1. Quintín Banderas
2. George Washington
3. Carlos Baliño
4. Panchito Gómez Toro
5. Héctor Rodríguez
6. Perucho Figueredo
7. Abel Santamaría
8. José María Pérez
9. Heriberto Duquesne
10. Ifraín Alfonso

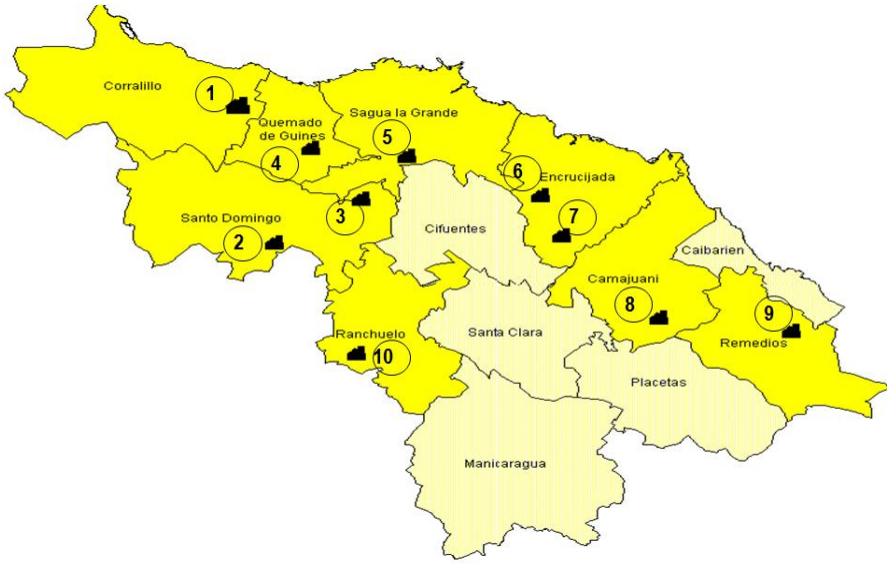


Figure 3-4. Agroindustrial Complex locations across the Villa Clara province.

3.2.2. Factors of variability in soil fertility parameters

Soil is never homogeneous. Heterogeneity exists even among soils that are classified as belonging to the same unit (Pleysier, 1995). Causes of spatial variability in soil fertility parameters include several factors. These factors occur as a result of the effect and interaction of various processes in the soil profile. Thus, the main consequences of spatial variability are related to the localized yield reduction, excessive fertiliser and water use, and nutrient losses (Parkin, 1993; Ehsani et al., 2009; Zaman et al., 2005; Shifteh et al., 2011; Mann et al., 2011). Therefore, many factors may cause variations among soils. According to Pleysier (1995) some of the more influential factors are:

1. Vegetation
2. Topography
3. Cropping and tillage practices

Vegetation. The type of plant life which covers a certain area of land can exert its influence in several ways. Changes in soil composition can occur when plant residue is incorporated into the soil. This is often the case after harvesting. Also, the structure of the soil is affected. The crop canopy and rooting system protect the soil from rain damage and severe temperatures, and make soil less susceptible to erosion and leaching.

Topography. Good top soil is often eroded from hilltops and ridges and can be found deposited in the lower areas. The slope angle and the slope length are important parts of topography, determining the amount of water that runs off or enters the soil.

Cropping and tillage practices. The distribution of soil nutrients, and thereby the composition of soils, varies among those that are heavily tilled and those that are worked only slightly or not at all. As a result, samples from fields with different cultivation practices are to be taken from different depths.

More than 80% of the areas of Cambisol and Vertisol is covered by sugarcane, the rest include other crops, native vegetation, pasture. The tillage practices are employed where sugarcane is cultivated (> 80%). Also, the 1.3% of the Cambisol areas is found with slightly undulating topography (2.1 – 4.1%), 15% show an undulating topography (4.1 – 8%), the rest tends to be flat. Generally, in Vertisol areas these slope patterns range from 1 to 2%.

The observed spatial variability in various soil properties that influence soil fertility will help farmers in making crop management decisions (Patil et al., 2011). Therefore, knowledge of spatial variability in soil fertility is important for site specific nutrient management (Adhikari et al., 2011; Patil et al., 2011; Shifteh et al., 2011).

One solution to this problem is variable-rate (VR) fertiliser application, which is more economical than uniform fertiliser application. For instance, Cox et al. (1999) reported an economic analysis of a case study involving sugarcane. The study showed a \$563 ha^{-1} benefit over five years when comparing VR technology with standard uniform input application. However, Mann et al. (2011) pointed out that VR fertiliser application depends on accurately identifying the underlying factors responsible for the variation in yield.

3.2.3. Soil group distribution

In Cuban agriculture and economy Cambisols (Brown soils) have a great importance, because they cover a large area, are well distributed all over the island, and are frequently used for sugarcane production. From the edaphic[2] point of view, they play an important role in understanding some of the pedogenetic processes that took place under tropical conditions, affected by climatic changes and a strong human influence (Villegas et al., 1995).

[2] **Edaphic** is a nature related to soil. Edaphic qualities may characterize the soil itself, including drainage, texture, or chemical properties.

Cambisols are soils at an early stage of soil formation. There is generally a brownish discoloration below the surface horizon to mark the beginning of pedogenesis. The horizon differentiation is weak. Cambisols are developed in medium and fine-textured materials derived from a wide range of rocks, mostly in alluvial, colluvial and aeolian deposits (FAO, 2001).

Cambisols are characterized by a cambic horizon occasionally associated with ochric, umbric, calcic or gypsic horizon. Because of their favourable aggregate structure and high content of weatherable minerals, they usually can be exploited for agriculture subject to the limitations of terrain and climate. Cambisols are present at all sampling locations and dominant at locations 5 to 9 on figure 3-5 the proportional distribution of these soils is illustrated. It can be seen that the largest proportions of Cambisol can be found in "Quintín Banderas" (46.12%), "José María Pérez" (44.80%) and "Ifraín Alfonso" (49%).

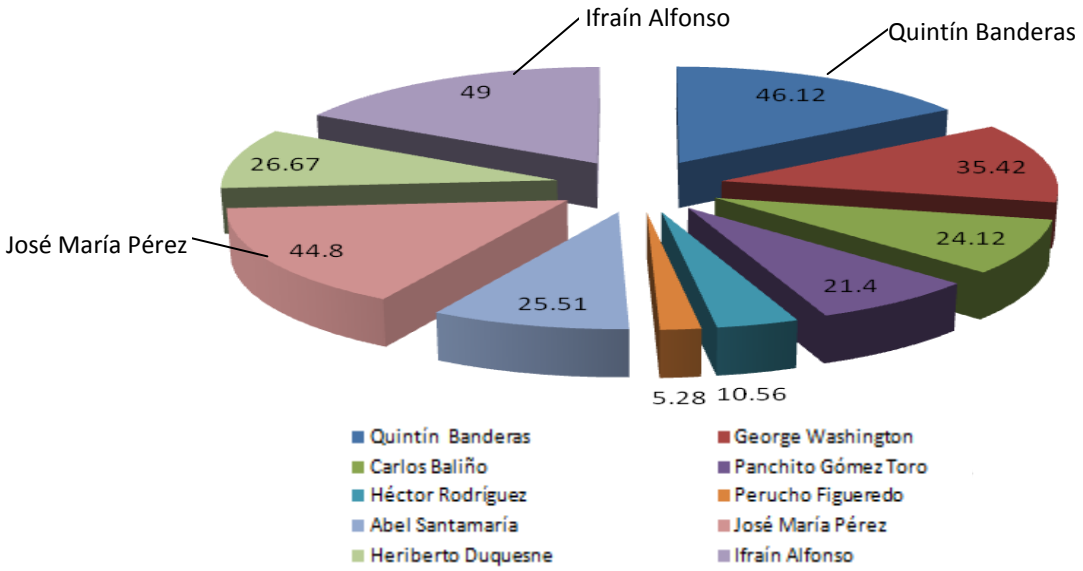


Figure 3-5. Percent distribution of Cambisol in CAI of Villa Clara province.

In Cuba, Vertisols are most wide-spread in the eastern part of the island. The mineralogical composition of the fraction of <1 μm includes several paragenetic associations[3]. The structural and functional role of clay (<1 μm) determines the specific properties of Vertisols, such as their high density, considerable swelling on moistening and shrinkage on drying, their high cation-exchange capacity, and others. When a

[3] **Paragenetic association** is the association and order of crystallization of minerals in a rock. Also, it is known as mineral sequence; paragenetic sequence.

Vertisol is used for agriculture the initial weakly water-resistant aggregates are destroyed and the clay fraction of the soil is peptized (Chizhikova, 2005).

This group of soils is represented in 8 of the 10 sampling locations. It is the most extended and occupies the first place in CAI "Abel Santamaría". It also has a very high significance and is ranged second in "Panchito Gómez Toro", "Héctor Rodríguez" and "Perucho Figueredo".

The proportional distribution of this soil group is illustrated in figure 3-6. The biggest values (>22%) correspond to "Abel Santamaría" (45.71%), "Perucho Figueredo" (32.73%) and "Héctor Rodríguez" (22.23%).

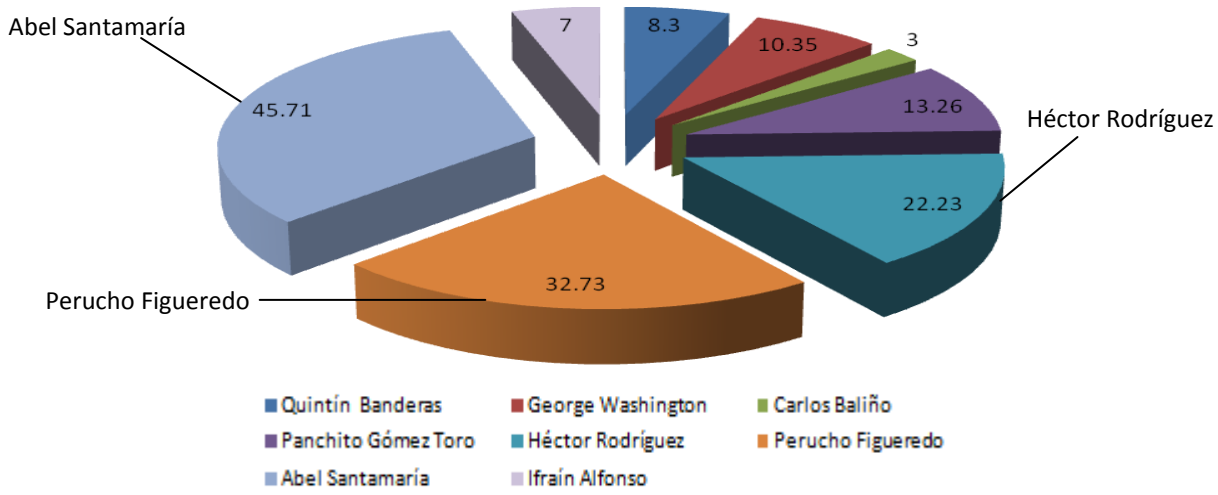


Figure 3-6. Percent distribution of Vertisol in CAI of Villa Clara province.

3.2.4. Soil sampling design for Vis/NIR calibration at landscape level

Landscape comprises the visible features of an area of land. According to Kizos and Koulouri (2005) agricultural landscapes are usually defined as the visual result of land uses and management systems in an area. The main criterion taken into account when selecting these landscape areas is the wide variation of edaphic factors linked to the soil fertility parameters concentration of Cambisol and Vertisol groups, as this variation is needed to build and evaluate good calibration models. The soil samples were collected on cloudless days in several fields from Cambisol and Vertisol groups of 10 CAI and within a field on Cambisol soil from an experimental area of the Central University "Marta Abreu" of Las Villas.

The samples were split into two datasets, one for calibration and the other for validation at landscape level. These samples were selected in 2007 between March and April on

189 and 144 fields for Cambisol and Vertisol respectively, ranged from 2.2 to 2.8 ha in both groups.

The soil samples were taken from the plough layer (0 – 20 cm) following the sampling scheme proposed by SERFE. In this system each soil sample consisted of 30 subsamples which were taken from subplots located across a diagonal line on the field, which starts and finishes 10 m from the field borders. The soil sampling was done before planting when the fields were ready to be furrowed. The samples were collected using a sampling auger with a footrest (Figure 3-7).

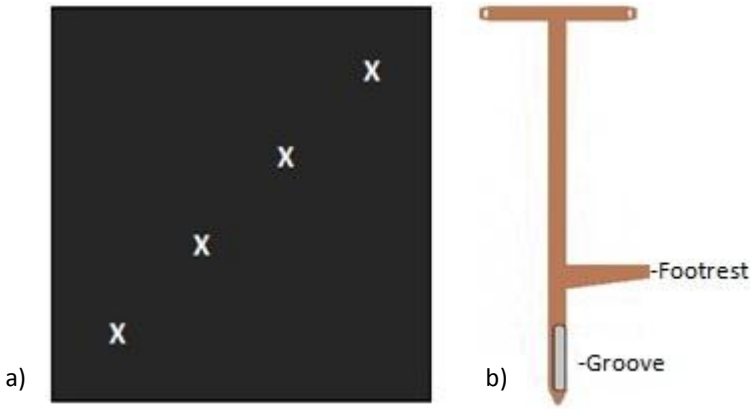


Figure 3-7. Soil sampling design for landscape (a) Scheme pattern, (b) Sampling auger.

Each sample was bagged separately into plastic bags with identification labels, air-dried at room temperature to constant weight, sieved with a 0.5 mm sieve and homogenized before chemical analysis.

3.2.5. Soil sampling design within a field

The field subsamples were selected from two fields of 2.03 ha for Cambisol (in June, 2012) and 1.96 ha for Vertisol (at the same period of those used for calibration). These fields are placed on Santa Clara and Sagüa la Grande municipalities, respectively.

These samples were used for establishing an independent validation set for assessing the soil fertility variation within a field. Also, the soil samples were collected at the same depth by using a stratified random sampling scheme (Figure 3-8). The fields were divided into several quadrants (strata) from which each soil core were selected separately and randomly. Each stratum was sampled in proportion to the total.

During the handling process the samples were bagged, labeled, air-dried at room temperature to constant weight, sieved with a 0.5 mm sieve, homogenized, and then measured with conventional chemical analyses.

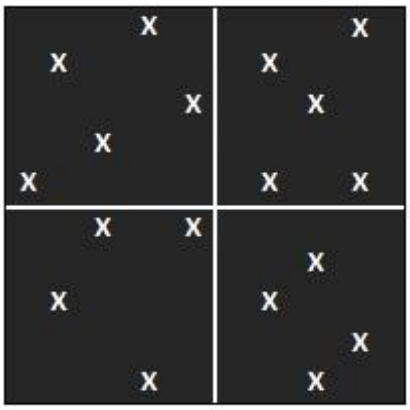


Figure 3-8. Soil sampling design pattern for assessing the variation within a field.

3.2.6. Chemical analyses

The collected soil samples were analysed for OM, K_2O and P. These soil fertility parameters are some of the most important for sugarcane growth. Also, these are the most common analysed parameters within the SERFE system. The chemical analyses were done in the analytical chemistry laboratories of the Territorial Station for Sugar Cane Research (ETICA, Spanish acronym) and in the Agricultural Research Centre (CIAP, Spanish acronym) belonging to the Central University "Marta Abreu" of Las Villas (applying the same methods used by SERFE). The following is a brief description of these conventional methods:

Walkley-Black method for organic matter

In this method, 20 ml of $K_2Cr_2O_7$ is added through a pipette to between 0.5 g and 1.0 g of soil placed in a 500 ml conical flask. The soil and dichromate were mixed by gently swirling the flask, followed by addition of 20 ml of concentrated H_2SO_4. The flask was again swirled gently to allow soil to have good contact with the reagent. The contents of the flask were allowed to stand for 30 minutes, followed by dilution with 200 ml of water (Sahrawat, 1982). The addition of H_3PO_4 to the digestive mix after the sample has cooled has been used to help eliminate interferences from the ferric (Fe^{3+}) iron that may be present in the sample. The chemistry of this extraction procedure (Schumacher, 2002) is as follows: $2Cr_2O_7^{2-} + 3\,C^0 + 16H^+ = 4Cr^{3+} + 3CO_2 + 8H_20$.

Olsen method for phosphorus

The method is based on the use of the HCO_3^-, CO_3^{2-} and OH^- in the pH 8.5, 0.5*M* $NaHCO_3$ solution to decrease the solution concentrations of soluble Ca_2 by precipitation as $CaCO_3$ and soluble Al^{3+} and Fe^{3+} by formation of Al and Fe oxyhydroxides, thus increasing P solubility. The increased surface negative charges and/or decreased number of positive sorption sites on Fe and Al oxide surfaces at high pH levels also enhance desorption of available P into solution. An Olsen P value of 10 mg P kg^{-1} is generally considered to be optimal for plant growth. This is lower than the critical values used for the Bray and Kurtz P-1, Mehlich-1 and Mehlich-3 soil tests because the Olsen extractant removes less P from most soils than the acidic extractants (Kovar and Pierzynski, 2009).

Oniani method for phosphorus and potassium

This method is commonly used in Cuba for determining phosphorus and potassium. It is based on the extraction of P with a concentrated solution of 0.1 *M* H_2SO_4 at a 1:25 soil: solution ratio and 3 minutes shaking time. Under these conditions it is possible to make vigorous extractions and the values obtained are very high. Four standard reagent solutions need to be prepared for the extractions and analytical reactions before setting up the colorimetric assay. These reagents are concentrated sulfuric acid (H_2SO_4), ammonium molybdate, 1-amino-2-naphthol-4-sulfonic acid, 4 *M* hydrochloric acid (4 *M* HCl), and 4 *M* Ammonium hydroxide (4 *M* NH_4OH). Approximately 15 ml of the P extractant is taken for analysing the K concentration, which is determined in a flame photometer.

The methods used for conventional chemical analyses are summarised in Table 3-1.

Table 3-1

Analytical methods for determination of soil fertility parameters.

Soil fertility parameter	Conventional method	Extraction	Determination
Organic matter	Walkley-Black	Digestion $K_2Cr_2O_7 + H_2SO_4$	Colorimetry
Available potassium	Oniani	0.1 *M* H_2SO_4	Flame photometry
Available phosphorus	Olsen	0.5 *M* $NaHCO_3$	Colorimetry
Available phosphorus	Oniani	0.1 *M* H_2SO_4	Colorimetry

The levels of the selected soil fertility parameters reported by different authors are summarised in Table 3-2. These ranges and levels were used for interpreting the concentrations measured for the studied soils.

Table 3-2

Classification of levels of the selected soil fertility parameters.

Soil fertility parameter	Level						Source
	Very low	low	Medium	Satisfactory	High	Very high	
OM	<1.5	1.5 -3.0	3.1-5.0	-	>5.0	-	López et al., 1981
K_2O	<6.2	≥6.2<8.8	≥8.8<13.8	-	≥13.8<32	≥32	Villegas et al., 1999
Olsen P	-	<1.14	1.14-2.29	2.29-4.12	>4.12	-	Jackson, 1964 Bingham, 1962 Hami, 1974
Oniani P	-	<6	6-11	-	11-15	>15	Villegas et al., 1999

OM in %; K_2O in mg K_2O 100 g^{-1} d.s[4], Olsen P and Oniani P in mg P 100 g^{-1} d.s

3.2.7. Data processing

The histograms of the average soil chemical properties per field were calculated in MATLAB 7.9 (R2009b, The Mathworks, Nattick, MT). Basic descriptive statistics (range, mean, median, skewness, kurtosis, minimum, maximum, standard deviation, coefficient of variation) and Pearson correlations were obtained by processing data with the statistical functions included in the Analyses Toolpack of Microsoft Excel 2007. Standard error of laboratory (SEL) was calculated as:

$$SEL = \sqrt{\frac{\sum_{i=1}^{N}[\sum_{j=1}^{R}(y_{ij}-\bar{y}_j)^2/(R-1)]}{N}} \qquad (3.2.1)$$

where y_{ij} is the *j*th replicate of the *i*th sample, $\bar{y}_j$ is the reference method mean value of all the replicates of the *i*th sample, *N* is the number of samples, and *R* is the number of replicates. In order to compare the variability of the soil fertility parameters among themselves across the landscape and within a field, the coefficient of variation (CV) was used. The results were categorized into the three classes proposed by Aweto (1982), where CV < 25% = low variability, 25 < CV < 50% = moderate variability, 50 < CV <100% = high variability. Factor analysis was used to group the four soil fertility parameters into factors based on the correlation matrix of the variables using the principal component analysis method of factor extraction in SPSS software. Principal component analysis was used as the method of factor extraction because it required no prior estimates of the amount of variation of each soil variable that would be explained by the factors (Brejda et al., 2000; Ayoubi et al., 2011).

[4] d.s- dry soil

3.3. Results and discussion

3.3.1. Frequency distribution of soil fertility parameters at landscape level

The histograms of the different soil fertility parameters at the landscape level, obtained by measuring an average value per field, are illustrated in Figure 3-9. The distribution for OM is weakly bimodal with two local maxima at the intervals from 2.61 to 2.73% and 2.86 to 2.98%. These intervals contain respectively 33 and 32 soil samples, which is respectively 17.46% and 16.93% of the analysed soil samples. According to the classification scale used in this research these values were evaluated as low. The histogram was positively skewed with a skewness coefficient of 0.76, which means that it was skewed towards the larger OM contents. The kurtosis coefficient was positive, which indicated a relatively peaked distribution.

Also, for K_2O content, the distribution was weakly bimodal, with two local maxima at the intervals comprised between 14.73 to 16.62 mg 100 g^{-1} d.s and 20.40 to 22.29 mg 100 g^{-1} d.s. These intervals respectively contain 13.76 and 11.64% of the analysed soil samples. These values are classified as high and the skewness and kurtosis coefficients obtained are positive.

For Olsen P, the distribution is positively skewed (skewness coefficient 1.12) as the tail from the central local maximum is longer towards the higher values. The distribution is also relatively peaked with a positive kurtosis coefficient (0.55). The local maximum interval ranges from 2.20 to 2.37 mg 100 g^{-1} d.s and contains 23.28% of the analysed soil samples. According to the classification scale for Olsen P used in this research, these interval values were classified between medium and satisfactory.

The distribution analyses for Oniani P content showed a histogram with one local maximum at the interval of 12.21 to 13.62 mg 100 g^{-1} d.s, corresponding to a frequency distribution of 30 soil samples which were equivalent to 15.87% of the total soil samples analysed. These values were classified as high, according to the scale used in this research. This histogram showed a skewness coefficient of 0.83, while the kurtosis coefficient was negative (-0.10).

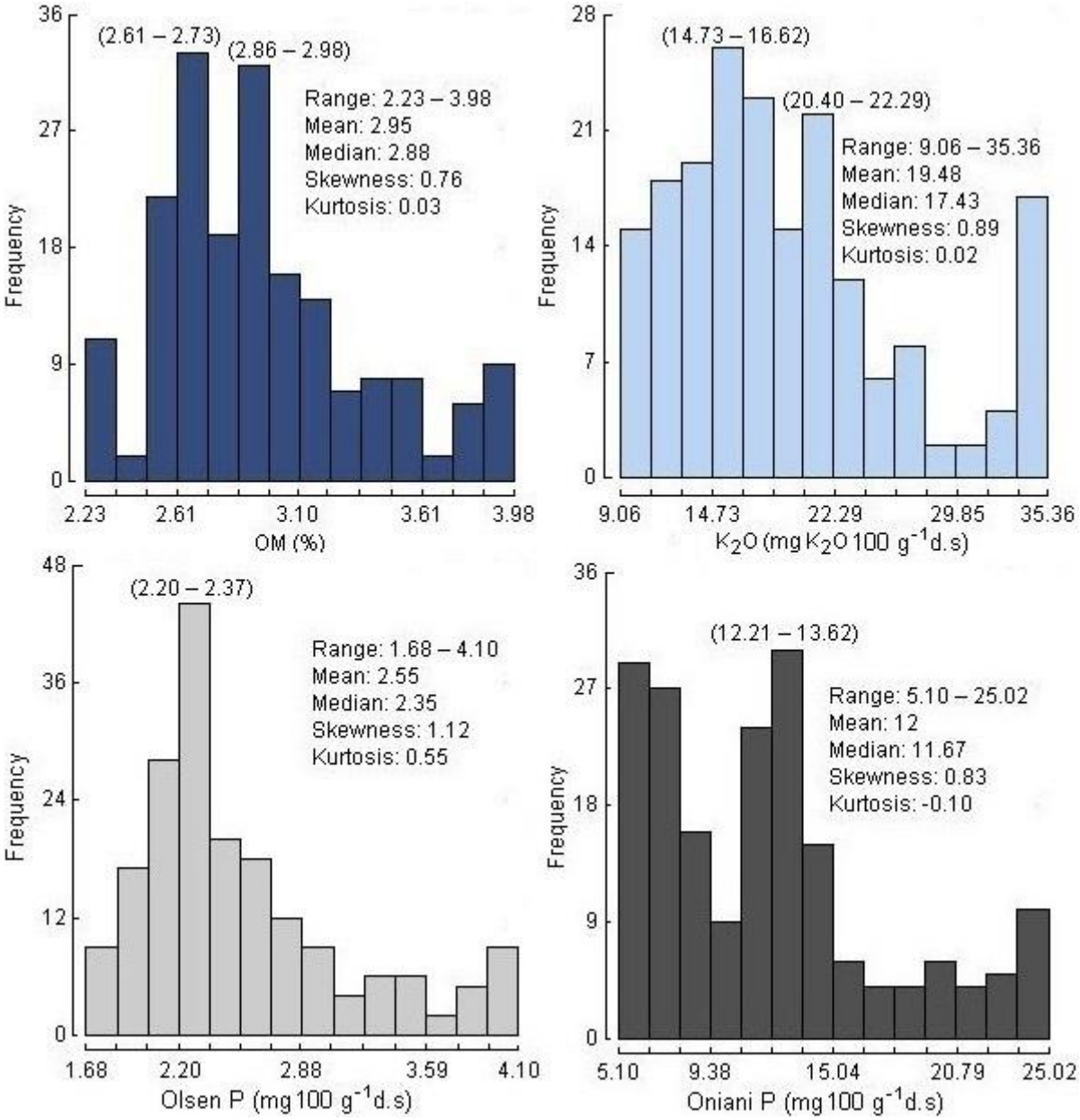


Figure 3-9. Frequency distribution of soil fertility parameters in Cambisol landscape.

The results of the frequency distribution of soil fertility parameters measured for Vertisols at the landscape level are presented in Figure 3-10. The distribution for OM content showed a histogram with one local maximum at the interval 2.55 – 2.71%. The values comprised at this interval corresponded to a frequency distribution of 26 soil samples. This number of samples was equivalent to 18.05% of the total soil samples analysed in this set, and the values were classified as low, according to the scale used. This histogram showed a positively skewed distribution with a skewness coefficient of 0.12 and a negative kurtosis coefficient (-0.64).

The bimodal distribution for K_2O content showed two local maxima at the intervals 6.62 – 9.15 mg 100 g^{-1} d.s and 11.60 – 13.95 mg 100 g^{-1} d.s. These intervals were represented by 22 and 26 soil samples, which were equivalent to 15.27 and 18.05% respectively of the total soil samples analysed. In the first interval these values were classified between low and medium and in the second interval were classified between medium and high. Potassium does not move readily in most soils; however it is more mobile than phosphorus (Hodges, 2010).

For Olsen P a right-skewed histogram was obtained with a positive skewness coefficient of 1.06 and a relatively peaked distribution where the kurtosis coefficient was positive (0.01). The interval of the local maximum was 1.32 – 1.77 mg 100 g^{-1} d.s. In this interval a frequency of 47 samples was observed, which represented 32.64% of the analysed soil samples. These values were classified as medium according to the classification scale used for Olsen P in this research.

The frequency distribution of the P content analysed by the Oniani method showed a histogram with one local maximum at the interval 4.41 – 6.25 mg 100 g^{-1} d.s. This interval corresponded to a frequency distribution of 39 soil samples equivalent to 27.08% of the analysed soil samples. These values were classified between low and medium, according to the scale used. Also, the histogram showed a positive skewness coefficient of 0.93, while the kurtosis coefficient was negative (-0.42).

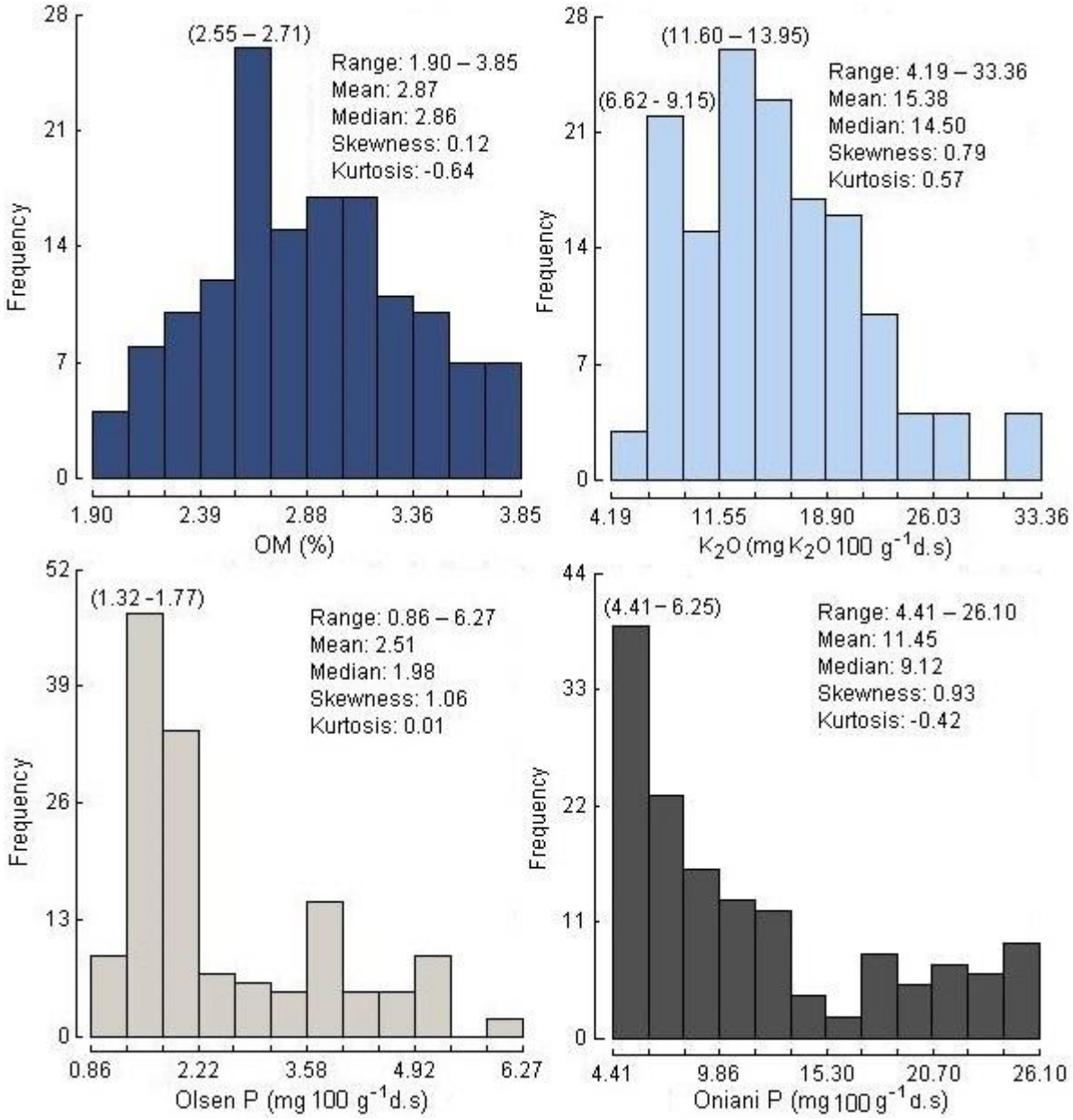


Figure 3-10. Frequency distribution of soil fertility parameters in Vertisol landscape.

3.3.2. Spatial variability of soil fertility parameters at landscape level

The descriptive statistics of the soil fertility parameters for the Cambisol soil samples from different fields in the Villa Clara province are listed in table 3. The higher coefficients of variation (CV) were observed in K_2O (36.49%) and Oniani P (44.62%). From the CV limits used in this research both parameters had a moderate variability (25 < CV < 50%) across the landscape. On the other hand the lowest CV was observed for OM (14.08%). This extensive degree of spatial variability provides the opportunity to apply site-specific fertiliser management strategies to reduce misapplications by improving the match between plant fertiliser requirements and fertiliser supply.

Table 3-3

Basic statistics for concentrations of soil fertility parameters in Cambisol landscape.

Soil fertility parameter	Mean	Min [a]	Max [b]	SD [c] (±)	CV [d] (%)
OM	2.95	2.23	3.98	0.42	14.08
K_2O	19.48	9.06	35.36	7.11	36.49
Olsen P	2.55	1.68	4.10	0.58	22.87
Oniani P	12.00	5.10	25.02	5.36	44.62

OM in %; K_2O in mg K_2O 100 g^{-1} d.s, Olsen P and Oniani P in mg P 100 g^{-1} d.s

[a] Min- minimum observed value; [b] Max- maximum observed value; [c] SD- standard deviation in the reference data; [d] CV- coefficient of variation

The spatial variability influences the size and number of soil samples needed to characterize the soil fertility parameters of an area of interest, which is in line with Corwin et al. (2003). The variability classified as moderate (K_2O, Oniani P) and low (OM, Olsen P) was probably due to fertiliser management practices applied to these areas. This soil fertility parameters variability can be a factor for the spatial variability in sugarcane yield reported[5] in these areas on Cambisol soils (Table 3-4). Therefore, quantifying the spatial variability of crop yield and soil fertility parameters can be important for decision making in site specific crop management. Consequently, detailed studies of spatial variations of soil fertility parameters and sugarcane yields are needed to support the application of precision agriculture in the Villa Clara province.

[5] Pérez, H., Rodríguez, I., Betancourt, Y., Gómez, J.M., Más, R., Rodríguez, L., Brito, G., Gutiérrez, J.L., 2001. Brief. Assessment of land physical condition in CAI from Villa Clara province (10 briefs). National Research Institute on Sugar Cane. La Habana.

Table 3-4

Distribution of Cambisol soils and spatial variability in sugarcane yield.

Municipality	CAI [a]	Area [b] (ha)	Yield				
			Mean	Min [c]	Max [d]	SD [e]	CV [f] (%)
			t ha^{-1}				
Corralillo	Q. B.	17 101.80	63.10	19.30	96	19.30	32.98
Santo Domingo	G. W.	14 762.60	51.65	18.70	105.50	18.70	50.21
Santo Domingo	C. B.	6 147.10	63.56	32.10	92	32.10	31.79
Quemado de Güines	P. G. T.	15 342.50	62.03	23.20	96	23.20	35.36
Sagua La Grande	H. R.	11 214.80	71.43	24	108	24	35.82
Encrucijada	P. F.	8 176.20	64.51	22.50	96	22.50	37.74
Encrucijada	A. S.	7 945.10	76.91	39.50	97.10	39.50	20.24
Camajuaní	J. M. P.	13 782.40	70.07	22.70	102.20	22.70	28.92
Remedios	H. D.	6 424.90	50.72	24	96	24	40.98
Ranchuelo	I. A.	6 403.60	52.60	25.15	77.99	25.15	24.57

Number of samples (n) = 115; [a] CAI- Agroindustrial complexes; [b]Area- Total area of each CAI;
[c] Min- minimum observed value; [e] Max- maximum observed value;
[f] SD- standard deviation in the reference data; [f] CV- coefficient of variation of yield in one year;
Q.B.- Quintín Banderas; G.W.- George Washington; C.B- Carlos Baliño; P.G.T.- Panchito Gómez Toro;
H.R.- Héctor Rodríguez; P.F.- Perucho Figueredo; A.S.- Abel Santamaría; J.M.P.- José María Pérez;
H.D.- Heriberto Duquesne; I.A.- Ifraín Alfonso.

These results could be a motivation towards acquiring knowledge related to the degree of spatial variability across this landscape (Table 3-5).

Table 3-5

Variation in the fertility parameters of Cambisol soils situated across the province.

CAI	Soil fertility parameters			
	OM	K_2O	Olsen P	Oniani P
Q.B.	medium	low	satisfactory	very high
G.W.	low	medium – high	medium	low – medium
C.B.	medium	medium	medium	medium
P.G.T.	medium	high – very high	medium – satisfactory	high – very high
H.R.	medium	low	medium	medium – high
P.F.	low	medium	medium	medium
A.S.	low	medium – high	medium	low – medium
J.M.	medium	high–very high	satisfactory	high – very high
H.D.	low – medium	medium	low – medium	low – medium
I.A.	low – medium	high – very high	low – medium	low – medium - high

OM in %; K_2O in mg K_2O 100 g^{-1} d.s, Olsen P and Oniani P in mg P 100 g^{-1} d.s
Q.B.- Quintín Banderas; G.W.- George Washington; C.B- Carlos Baliño; P.G.T.- Panchito Gómez Toro;
H.R.- Héctor Rodríguez; P.F.- Perucho Figueredo; A.S.- Abel Santamaría; J.M.P.- José María Pérez;
H.D.- Heriberto Duquesne; I.A.- Ifraín Alfonso.

In Table 3-6 the descriptive statistics of the results obtained by using conventional chemical analyses on the soil samples from Vertisol landscape are presented. In this case phosphorus was the soil fertility parameter with a higher variability than the others, independently of the chemical method used for its analysis. The higher value was observed in Oniani P with 57.17%, followed by Olsen P with 50.89%. Both values accounted for a high variability of these parameters.

Table 3-6

Basic statistics for concentrations of soil fertility parameters in Vertisol landscape.

Soil fertility parameter	Mean	Min [a]	Max [b]	SD [c] (±)	CV [d] (%)
OM	2.87	1.90	3.85	0.46	16.01
K_2O	15.38	4.19	33.36	6.04	39.28
Olsen P	2.51	0.86	6.27	1.28	50.89
Oniani P	11.45	4.41	26.10	6.54	57.17

OM in %; K_2O in mg K_2O 100 g^{-1} d.s, Olsen P and Oniani P in mg P 100 g^{-1} d.s

[a] Min- minimum observed value; [b] Max- maximum observed value; [c] SD- standard deviation in the reference data; [d] CV- coefficient of variation

The high coefficients of variation observed in soil P levels are dependent on management practices. It seems that the applications of P fertilisers increased the variability of this nutrien in soil. Continued applications of P fertiliser are often required to maintain a given level of crop production. Then, the magnitude of P fertilisation has been constantly increased and applied in a uniform way across these fields. However, crops usually take up approximately 10 - 15% of the added P fertiliser, while the remainder is accumulated in the soil. Consequently, spatial variations of this soil fertility parameter can appear both horizontally and vertically. The lower CV corresponded to the OM, just like in Cambisol landscape. This parameter showed a low variability across this area, while the variability of K_2O was classified as moderately variability. These results indicated a considerable variability in soil fertility parameters in Vertisol landscape, particularly for Olsen P and Oniani P. Fertiliser recommendations are commonly targeted for an average soil and management system, and then are applied for general soil types across a whole province. Thus, VR fertiliser application for different fields, hence, should be considered as an important method for making soil fertility distribution more uniform. Understanding the magnitude and pattern in spatial variability of these soil fertility parameters and sugarcane yield is necessary for improved management options relating to application of fertiliser in areas on Vertisol

soil (Table 3-7). The spatial variability of soil fertility parameters and sugarcane yield is a critical factor for making VR inputs of fertilisers. In comparison, there were small differences in CV between reported yields in Cambisol and Vertisol areas.

Table 3-7

Distribution of Vertisol soils and spatial variability in sugarcane yield.

Municipality	CAI [a]	Area [b] (ha)	Yield				
			Mean	Min [c]	Max [d]	SD [e]	CV [f] (%)
			t ha^{-1}				
Corralillo	Q. B.	17 101.80	32.98	13.30	61.60	14.38	43.59
Santo Domingo	G. W.	14 762.60	36.47	18.20	66	8.92	24.46
Santo Domingo	C. B.	6 147.10	27.66	19	29.10	3.18	11.48
Quemado de Güines	P. G. T.	15 342.50	27.80	16	36.40	4.93	17.72
Sagua La Grande	H. R.	11 214.80	32.89	18.80	52	6.74	20.49
Encrucijada	P. F.	8 176.20	29.57	27.50	42.50	7.87	26.61
Encrucijada	A. S.	7 945.10	34.19	24.40	61.10	9.77	28.58
Ranchuelo	I. A.	6 403.60	38.94	31.92	43.35	4.87	12.50

Number of samples (n) = 92; [a] CAI- Agroindustrial complexes; [b]Area- Total area of each CAI;
[c] Min- minimum observed value; [e] Max- maximum observed value;
[f] SD- standard deviation in the reference data; [f] CV- coefficient of variation of yield in one year;
Q.B.- Quintín Banderas; G.W.- George Washington; C.B- Carlos Baliño; P.G.T.- Panchito Gómez Toro;
H.R.- Héctor Rodríguez; P.F.- Perucho Figueredo; A.S.- Abel Santamaría; I.A.- Ifraín Alfonso.

In table 3-8 the different levels of these soil fertility parameters, tested for this research throughout the municipalities and CAI from Villa Clara province, are shown.

Table 3-8

Variation in the fertility parameters of Vertisol soils distributed across the province.

CAI	Soil fertility parameters			
	OM	K_2O	Olsen P	Oniani P
Q. B.	low	medium	medium	medium
G. W.	medium	very low - low	low - medium	low - medium
C. B.	medium	medium	low - medium	medium
P. G. T.	low	low - medium	medium	low - medium
H. R.	low - medium	medium - high	medium - satisfactory	medium - high
P. F.	low - medium	medium	medium	medium
A. S.	low	medium	low - medium	low – medium - high
I. A.	medium	high	satisfactory - high	high – very high

OM in %; K_2O in mg K_2O 100 g^{-1} d.s, Olsen P and Oniani P in mg P 100 g^{-1} d.s
Q.B.- Quintín Banderas; G.W.- George Washington; C.B- Carlos Baliño; P.G.T.- Panchito Gómez Toro;
H.R.- Héctor Rodríguez; P.F.- Perucho Figueredo; A.S.- Abel Santamaría; I.A.- Ifraín Alfonso.

3.3.3. Factor analysis at landscape level

For the four soil fertility parameters measured, a maximum of four factors explain the total variance of each factor. An eigenvalue analysis allows the identification of the significant factors that collectively represent the major proportions of the total variability. The first factor shows an eigenvalue more than 1, for that reason is the most significant for explaining the system variance compared to the remaining factors (Table 3-9). When an eigenvalue is less than 1 the factor explains less variance than the individual attribute. This is in line with Shukla et al. (2006) and Ayoubi et al. (2011).

Table 3-9

Initial eigenvalues, proportion of variance and cumulative variance in Cambisol landscape.

Factor	Eigenvalue	Proportion of variance	Cumulative variance
		%	
1	3.42	85.55	85.55
2	0.34	8.41	93.96
3	0.15	3.84	97.80
4	0.09	2.20	100

The rotated factor matrix shows that Factor 1 explained 57.71% of the total variance respectively, with positive loadings from all the soil fertility parameters. In general, the factor loadings showed that no particular soil fertility parameter evaluated had a much more pronounced influence than other. All contributed quite equally to the same factor that produces the variability in the study (Table 3-10).

Table 3-10

Varimax rotation and comunalities for soil fertility parameters in Cambisol landscape.

Soil fertility parameter	Factor matrix	Rotated factor matrix	Comunality estimates
OM	0.964	0.828	0.94
K_2O	0.868	0.401	0.99
Olsen P	0.942	0.792	0.89
Oniani P	0.923	0.914	0.94
Eigenvalue	3.42	2.31	-
Variance %	85.55	57.71	-
Cumulative variance %	85.55	57.71	-

Also, the importance of each soil fertility parameter, as to its contribution to all of the factors, is considered by its communality value. This factor explained less than 90% of variance in Olsen P; 99% of variance in K_2O content, following of OM and Oniani P with 94%.

In Vertisol landscape only the Factor 1 shows an eigenvalue > 1, and then it is the most significant factor for explaining the system variance (Table 3-11).

Table 3-11

Initial eigenvalues, proportion of variance and cumulative variance in Cambisol landscape.

Factor	Eigenvalue	Proportion of variance	Cumulative variance
		%	
1	3.47	86.64	86.64
2	0.33	8.12	94.76
3	0.18	4.42	99.18
4	0.03	0.82	100

The Factor 1 accounted for an 86.64% of the total variance with high positive loadings from all the soil fertility parameters. This factor accounted for 49.17% of the total variance when rotated, also with positive loadings from all the soil fertility parameters. According to the factor loadings all the soil fertility parameters contributed quite equally to the same factor that produces the variability. The communality estimates explained 89% of variance in OM and more than 90% in the remaining factors (Table 3-12).

Table 3-12

Varimax rotation and comunalities for soil fertility parameters in Vertisol landscape.

Soil fertility parameter	Factor matrix	Rotated factor matrix	Comunality estimates
OM	0.921	0.522	0.89
K_2O	0.897	0.399	0.94
Olsen P	0.963	0.856	0.98
Oniani P	0.940	0.896	0.98
Eigenvalue	3.47	1.97	-
Variance %	86.64	49.17	-
Cumulative variance %	86.64	49.17	-

3.3.4. Correlation between soil fertility parameters at landscape level

Consistent and positive correlations were found between all the soil fertility parameters. Another significant aspect was that the higher correlation coefficients observed for K_2O, Olsen P and Oniani P were related to OM (Table 3-13) with the highest correlation coefficients between OM and phosphorus, at 0.89 for Olsen P and 0.88 for Oniani P.

Table 3-13

Pearson correlation coefficients among soil fertility parameters in Cambisol.

	OM	K_2O	Olsen P	Oniani P
OM	1.00			
K_2O	0.78*	1.00		
Olsen P	0.89*	0.68*	1.00	
Oniani P	0.88*	0.76*	0.83*	1.00

*Significant at the level of <0.001.

In agreement with these results Whitbread (1995) explained that it is generally accepted that soil organic matter has beneficial effects on soil fertility parameters, which in turn influences the productive capacity of soils. It is also accepted that OM is a major contributor of P as well as other nutrients to plants. The lowest correlation of 0.68 was between Olsen P and K_2O at 0.68, which is still significant. The correlation results for the Vertisol soil at landscape level showed that K_2O, Olsen P and Oniani P were strongly correlated with OM as for the Cambisol soils (Table 3-14).

Table 3-14

Pearson correlation coefficients among soil fertility parameters in Vertisol.

	OM	K_2O	Olsen P	Oniani P
OM	1.00			
K_2O	0.82*	1.00		
Olsen P	0.78*	0.75*	1.00	
Oniani P	0.83*	0.78*	0.96*	1.00

*Significant at the level of <0.001.

As a direct source of P, OM is one of the most important elements of the soil to consider in studying how P behaves in the agricultural fields across this province. The quantity of OM present in these agricultural fields resulted from the addition or loss of OM through decomposition. As van Schöll and Nieuwenhuis (2004) pointed out, a positive balance of OM is difficult to achieve. This means that if large amounts of OM are lost, it is difficult to restore its level in the soil. Even in favourable conditions and with good crop

management, this can take several decades, especially if during that time crops are grown that are almost completely removed with the harvest.

3.3.5. Frequency distribution of soil fertility parameters within a field

The histograms for soil fertility parameters analysed for the different subsamples taken from one Cambisol field are illustrated in Figure 3-11. For OM content a maximum of 9 soil samples was observed in the interval from 3.42 to 3.56% OM, which includes the largest OM value. This level of OM was classified as medium, according to the scale used in this research.

On the other hand, two intervals with 8 samples were observed, comprised between 2.54 to 2.69% and 2.69 to 2.83%, which were classified as low. The skewness coefficient was positive and near 0 (0.08), while the kurtosis coefficient was negative (-1.67) indicating a relatively flat distribution. These results are influenced by the effect of the agricultural practices of the farming systems used in Cuba for sugarcane production, and more specifically the soil fertilisation management. For instance, the redistribution of soil OM between particle-size fractions is most probably affected by land use.

The previous statement was based on the observation of Caravaca et al. (1999) related to the results obtained in their study, where total OM content of cultivated soils was higher in the -2 mm fraction, while in forest soils it was higher in the 2 - 20 mm fraction. Another example of the direct influence of farming systems on the soil OM content related to the mechanization patterns in the conventional Cuban farming systems. The use of the disc harrow causes soil disintegration and accelerates the OM oxidation. Agricultural practices with intensive machinery use and low OM input, according to Barzegar et al. (2000) cause deterioration of the soil structure and increase compaction.

The histogram with the frequency distribution of K_2O content is bimodal with two local maxima at the intervals comprised between 11.10 to 14.35 mg 100 g^{-1} d.s and 17.68 to 20.86 mg 100 g^{-1} d.s. These ranges represent respectively 24.32% and 21.62% of the total analysed soil samples. The first interval is classified as medium to high, while the second interval is classified as high. The skewness coefficient is positive (0.70) while the kurtosis coefficient is negative (-0.24).

For Olsen P the histogram is also bimodal with two local maxima at the intervals 2.15 - 2.35 mg 100 g^{-1} d.s and 2.75 - 2.95 mg 100 g^{-1} d.s. These local maxima obtained corresponded to a frequency distribution of 9 and 6 soil samples, which respectively, represent the 24.32% and 16.21% of the total soil samples analysed. The values comprised in the first interval were classified as medium and satisfactory and those of the second interval as satisfactory. The skewness coefficient was positive (0.44) while the kurtosis coefficient was negative (-0.76).

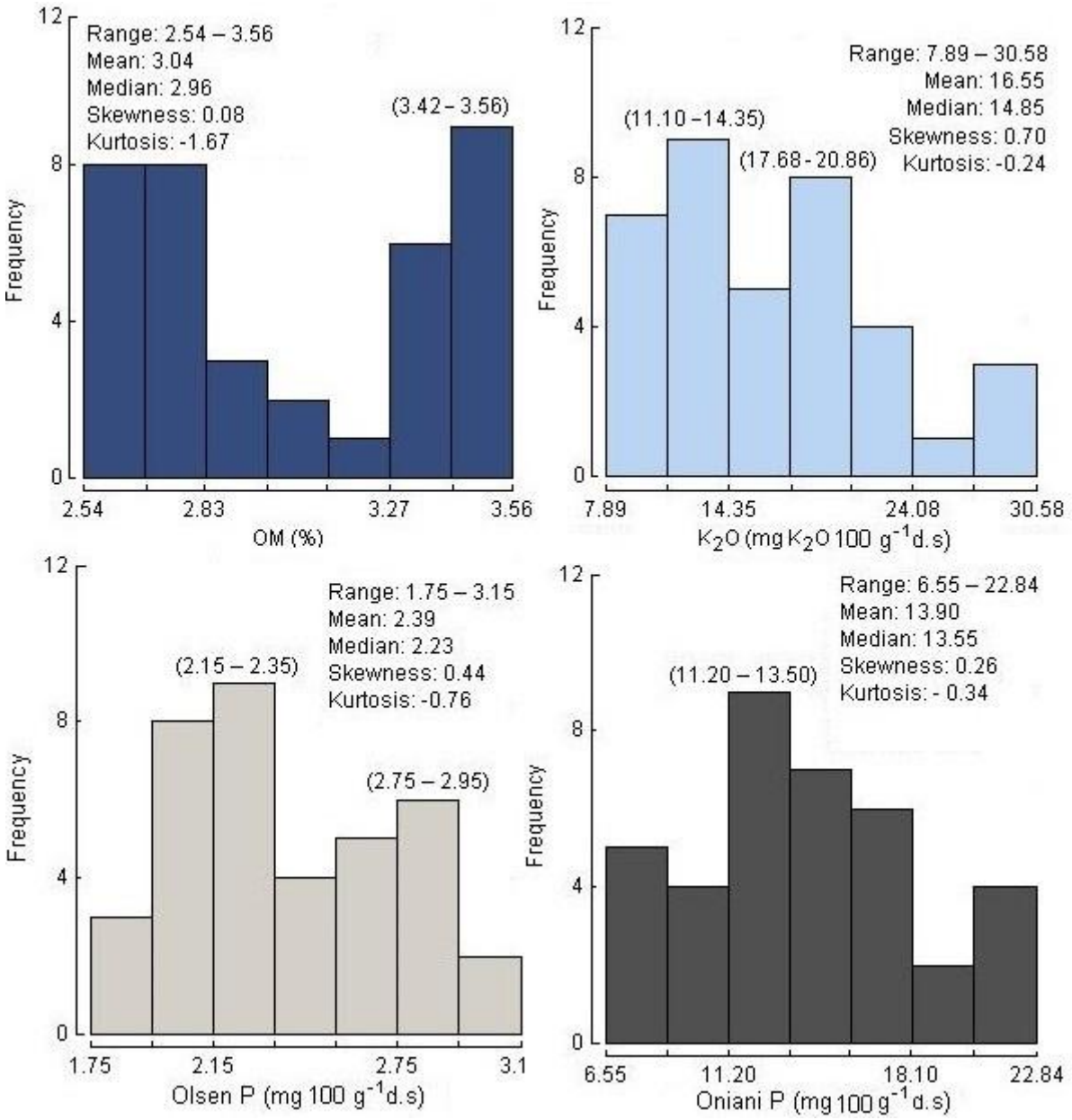


Figure 3-11. Frequency distribution of soil fertility parameters within a field on Cambisol soil.

In the histogram with the frequency distribution of the P content analysed by the Oniani method, a local maximum was observed at the interval of 11.20 to 13.50 mg 100 g^{-1} d.s, where 9 soil samples were included, which corresponded to the 24.32% of the total soil samples analysed (Figure 3-10). These values were classified between medium and high and in the second interval were only high. The skewness coefficient was positive (0.26) while the kurtosis coefficient was negative (-0.34).

In the histogram with the frequency distribution of the P content analysed by the Oniani method, a local maximum was observed at the interval of 11.20 to 13.50 mg 100 g^{-1} d.s, where 9 soil samples were included, which corresponded to the 24.32% of the total soil samples analysed (Figure 3-10). These values were classified between medium and high and in the second interval were only high. The skewness coefficient was positive (0.26) while the kurtosis coefficient was negative (-0.34).

The frequency distribution histograms for the subsamples taken from the same Vertisol field are given in Figure 3-12. In this sense a local maximum obtained for OM corresponded to a frequency distribution of 8 samples which represented 27.59% in the interval 2.99 – 3.23%. These values were classified between low and medium. The higher number of samples was towards the left side of this local maximum, and included a total of 17 or 58.62% of all those analysed in this set. Also, two intervals with the same frequency of 5 soil samples were observed. These two intervals comprised values between 2.50 to 2.75% and 2.75 to 2.99%, which were evaluated as low in both cases. The skewness and kurtosis coefficients were negative with values of -0.29 and -0.65 respectively.

For K_2O two local maxima with 8 soil samples in each one were observed at the intervals 11.47 – 14.50 mg 100 g^{-1} d.s and 14.50 – 17.57 mg 100 g^{-1} d.s. These 8 soil samples represented 27.59% of the total analysed in this set. In the first interval these values were classified among medium and high and in the second interval were high. The skewness coefficient was positive (0.03) while the kurtosis coefficient was negative (-0.35). The frequency distribution histogram for Olsen P shows a local maximum at the interval of 1.45 – 2.07 mg 100 g^{-1} d.s. The local maximum obtained corresponds to a frequency distribution of 13 soil samples which represent the 44.83% of the total soil samples analysed in this set. These values were classified as medium. Also, two intervals with the same number of soil samples (5) were observed at both sides (right and left) of the local maximum. These two intervals comprised values

between 0.83 to 1.45 mg 100 g^{-1} d.s and 2.07 to 2.69 mg 100 g^{-1} d.s respectively. The skewness and the kurtosis coefficients were positive with values of 1.10 and 0.18 respectively. In the histogram with the frequency distribution of Oniani P content, two local maxima can be observed at the intervals 4.46 – 7.25 mg 100 g^{-1} d.s and 9.95 – 12.80 mg 100 g^{-1} d.s. In the first interval 12 soil samples (41.38%) are included while in the second interval 6 soil samples (20.69%) are included. These values in the first interval were classified between low and medium and in the second interval between medium and high. The skewness and the kurtosis coefficients were positive with values of 1.06 and 0.51 respectively.

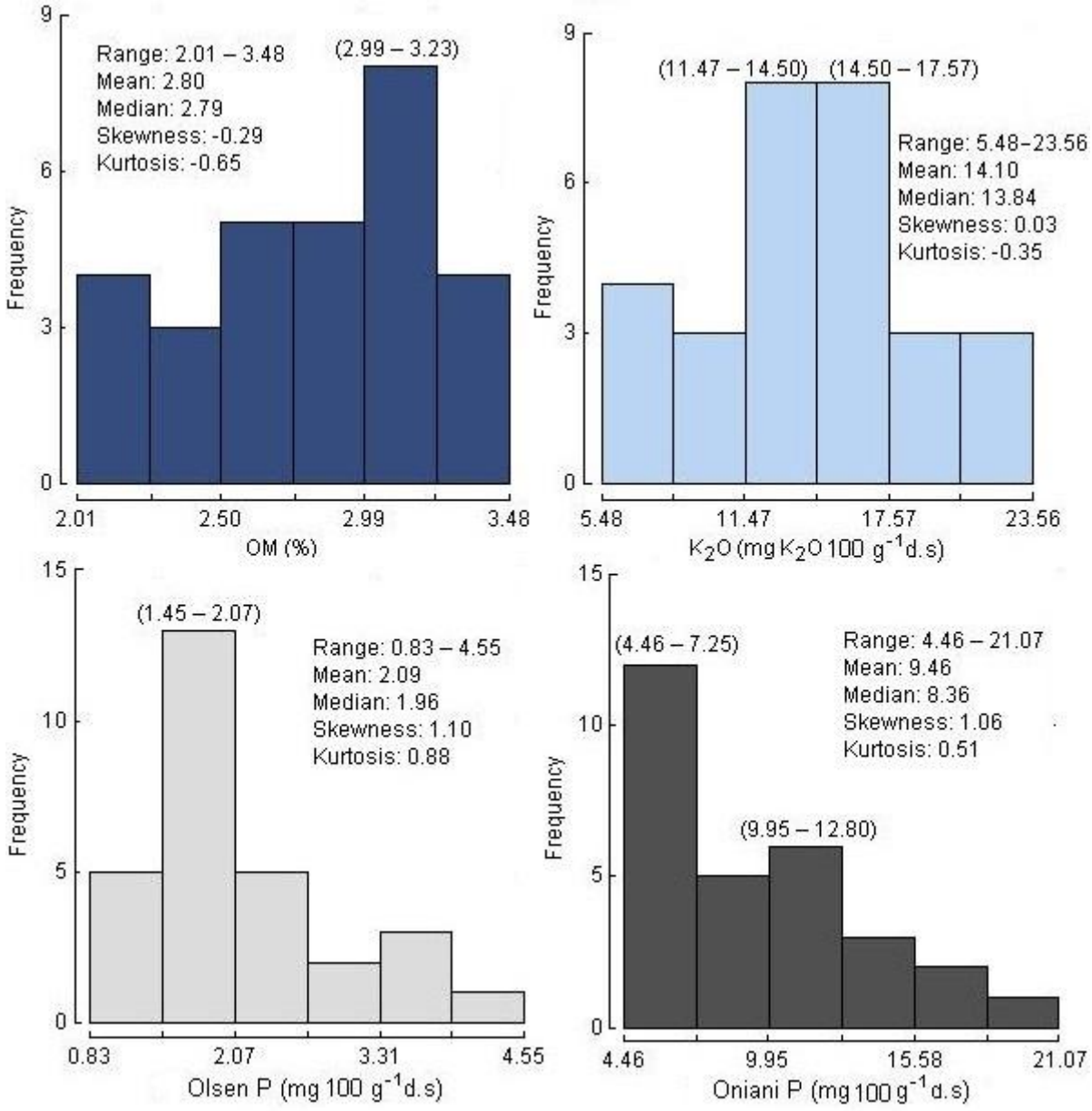


Figure 3-12. Frequency distribution of soil fertility parameters within a field on Vertisol soil.

3.3.6. Spatial variability of soil fertility parameters within a field

Table 3-15 shows the results of soil fertility parameters analysed within a field on Cambisol soil. According to the CV limits used in this research, none of these parameters had a high variability (50 < CV <100%). However, K_2O and Oniani P showed a moderate variability with 37.68% and 31.70% respectively. On the other hand the CV values for OM and Olsen P were less than 16%, which indicated a low variability for these parameters.

In Cambisol landscape, a higher variability was found for almost all soil fertility parameters, except for K_2O which at landscape level (36.49%) was lower than within a field (37.68%). However, the difference between K_2O determined at landscape level and K_2O obtained at field scale was smaller than those observed for the other soil parameters. Also, this parameter showed the highest value of CV within a field.

Table 3-15

Basic statistics for concentrations of soil fertility parameters within a field on Cambisol soil.

Soil fertility parameter	Mean	Min [a]	Max [b]	SD [c] (±)	CV [d] (%)
OM	3.04	2.54	3.56	0.37	12.16
K_2O	16.55	7.89	30.58	6.84	37.68
Olsen P	2.39	1.75	3.15	0.36	15.27
Oniani P	13.90	6.55	22.84	4.39	31.70

OM in %; K_2O in mg K_2O 100 g^{-1} d.s, Olsen P and Oniani P in mg P 100 g^{-1} d.s
[a] Min- minimum observed value; [b] Max- maximum observed value; [c] SD- standard deviation in the reference data; [d] CV- coefficient of variation

The basic statistics for soil fertility parameters of a Vertisol field are shown in Table 3-16. The CV results confirmed a low variability for OM (14.41%) and a moderate variability for K_2O, Olsen P and Oniani P.

Table 3-16

Basic statistics for concentrations of soil fertility parameters within a field on Vertisol soil.

Soil fertility parameter	Mean	Min [a]	Max [b]	SD [c] (±)	CV [d] (%)
OM	2.80	2.01	3.48	0.40	14.41
K_2O	14.10	5.48	23.56	4.59	32.57
Olsen P	2.09	0.83	4.55	0.92	43.95
Oniani P	9.46	4.46	21.07	4.45	47.02

OM in %; K_2O in mg K_2O 100 g^{-1} d.s, Olsen P and Oniani P in mg P 100 g^{-1} d.s
[a] Min- minimum observed value; [b] Max- maximum observed value; [c] SD- standard deviation in the reference data; [d] CV- coefficient of variation

In the Vertisol landscape, the CV values were higher for all the soil fertility parameters. However, for OM and K_2O there was no difference as to the variability classification. In both cases, landscape level and field scale, the variability was classified as low (OM) and moderate (K_2O). On the other hand, across the Vertisol landscape the CV values indicated a high variability for phosphorus (Olsen & Oniani) as previously discussed. This is in line with Weindorf and Zhu (2010), which found CV values of 49.73% for phosphorus and 29.72% for potassium.

3.3.7. Factor analysis within a field

The Factor 1 is retained according to the eigenvalue > 1 within a field on Cambisol soil. This eigenvalue implies the most significant importance for explaining the system variance. Also, this factor accounted for an 82.67% of the total variance (Table 3-17).

Table 3-17

Initial eigenvalues, proportion of variance and cumulative variance within a field on Cambisol soil.

Factor	Eigenvalue	Proportion of variance	Cumulative variance
		%	
1	3.31	82.67	82.67
2	0.36	9.01	91.68
3	0.21	5.33	97.01
4	0.12	2.99	100

After varimax rotation this factor accounted for 54.80% of the total variance, also with positive loadings from all the soil fertility parameters. The loading from K_2O was lower than the remaining parameters (Table 3-18).

Table 3-18

Varimax rotation and comunalities for soil fertility parameters within a field on Cambisol soil.

Soil fertility parameter	Factor matrix	Rotated factor matrix	Comunality estimates
OM	0.953	0.794	0.91
K_2O	0.861	0.377	0.98
Olsen P	0.918	0.762	0.85
Oniani P	0.903	0.917	0.93
Eigenvalue	3.31	2.19	-
Variance %	82.67	54.80	-
Cumulative variance %	82.67	54.80	-

However, according to the factor loadings the soil fertility parameters contributed in the same way to the same factor, which produces the variability. The communalies explained 85% of variance in Olsen P and more than 90% in the remaining factors.

The Factor 1 shows an eigenvalue > 1 within a field on Vertisol soil, and then it is retained as the most important factor for explaining the system variance. This factor accounted for an 84.74% of the total variance (Table 3-19).

Table 3-19

Initial eigenvalues, proportion of variance and cumulative variance within a field on Vertisol soil.

Factor	Eigenvalue	Proportion of variance	Cumulative variance
		%	
1	3.39	84.74	84.74
2	0.35	8.66	93.40
3	0.22	5.44	98.84
4	0.05	1.16	100

This factor accounted for 47.81% of the total variance when rotated, with positive loadings. The loadings for OM and K_2O were lower than those from the remaining soil fertility parameters. According to the factor loadings all the soil fertility parameters contributed in the same way to the same factor that produces the variability. The communality estimates explained 89% of variance in OM and more than 90% in the remaining factors (Table 3-12).

Table 3-20

Varimax rotation and comunalities for soil fertility parameters within a field on Vertisol soil.

Soil fertility parameter	Factor matrix	Rotated factor matrix	Comunality estimates
OM	0.899	0.447	0.89
K_2O	0.893	0.419	0.90
Olsen P	0. 950	0.866	0.97
Oniani P	0.938	0.887	0.97
Eigenvalue	3.39	1.91	-
Variance %	84.74	47.81	-
Cumulative variance %	84.74	47.81	-

3.3.8. Correlation between soil fertility parameters within a field

The results of the correlation analyses presented in Table 3-21 indicate that all the measured soil fertility parameters within a field on Cambisol soil exhibited positive correlations of diverse magnitude. These correlations were very similar to those obtained for the Cambisol soil samples at the landscape level. The highest correlation coefficients observed for K_2O, Olsen P and Oniani P were related to OM. According to Hodges (2010) higher OM levels can help reduce P fixation reactions, by binding Al, Fe and Ca, and forming soluble complexes with P. Another similarity to the landscape correlations was that the highest correlation coefficient was observed between OM and Olsen P at 0.85. Again, the lowest value was observed between Olsen P and K_2O at 0.65.

Table 3-21

Pearson correlation coefficients among soil fertility parameters within a field on Cambisol soil.

	OM	K_2O	Olsen P	Oniani P
OM	1.00			
K_2O	0.77*	1.00		
Olsen P	0.85*	0.65*	1.00	
Oniani P	0.83*	0.72*	0.78*	1.00

*Significant at the level of <0.001.

The correlation coefficients within a field on Vertisol soil are shown in Table 3-22. The correlation coefficients obtained between these parameters were very similar to those at landscape level. The higher values were observed at landscape level, except for the correlation coefficient between Olsen P and K_2O which was very similar (0.75 and 0.76 at landscape level and within a field respectively).

Table 3-22

Pearson correlation coefficients among soil fertility parameters within a field on Vertisol soil.

	OM	K_2O	Olsen P	Oniani P
OM	1.00			
K_2O	0.79*	1.00		
Olsen P	0.74*	0.76*	1.00	
Oniani P	0.79*	0.75*	0.95*	1.00

*Significant at the level of <0.001.

A higher correlation coefficient was observed between K_2O and OM for the Vertisol landscape. The best relationship (r= 0.95) was obtained between Olsen P and Oniani P. Bogrekci and Lee (2005a) also obtained a high correlation (r= 0.98) between total P and

Mehlich-1 P method, and stated that since the relationship is so strong, both methods can be used interchangeably for soils when cost of chemical analyses is an issue. A similar high correlation was also observed at landscape level (r= 0.96). On the other hand, lower value of the correlation coefficient of 0.74, which was still significant, was observed between OM and Olsen P.

3.4. Conclusions

Frequency distribution and ranges of soil fertility parameters for Cambisol and Vertisol soils at landscape level and within a field have been investigated by means of soil sampling and chemical analysis. In terms of measured nutrients, similarities were found between Cambisol and Vertisol when comparing the values observed on the local maxima of each frequency distribution histogram at landscape level and within a field. These results indicated that in several locations the soil fertility parameters analysed do not specify P or K_2O as a limiting soil nutrient, according to the classification scale used. For that reason the soil fertility, in terms of K_2O and P content, is not low for sugarcane production. However, essential differences were observed when comparing the spatial variability of these soil fertility parameters in each soil group. The high coefficients of variation (CV) obtained in this research might indicate a large spatial variability of the soil fertility parameters, mainly P and K_2O, for different fields within the province and even within a field. It means that soil analyses prior to fertilisation are needed to allow efficient fertilisation. The soil fertility parameters, P and K_2O, had a CV greater than 30% for both sets of soil groups demonstrating a high variability within these datasets. Only for organic matter was the variation in the present study considerably lower with a CV of 15%. This demands for a strategy to asses this variation in the soil fertility parameters within a field prior to VR fertilisation. In factor analysis the significance of the eigenvalues was used as a criterion for understanding the relationship between soil fertility parameters and factors. In all cases, soil fertility parameters were assigned to the Factor 1 for which their eigenvalues were the highest. Also, significant correlations were observed between all the investigated soil fertility parameters indicating that the other soil fertility parameters are highly correlated to the organic matter content both in Vertisols and Cambisols at the landscape level and within a field.

Chapter 4

Vis/NIR soil spectra analyses

4.1. Introduction

Vis/NIR spectroscopy is an efficient tool for analysing soil fertility parameters as mentioned in Chapter 2. Basically Cécillon et al. (2009) distinguished three types of Vis/NIR measurements for soils: (i) laboratory measurements, (ii) proximal sensing measurements and (iii) remote sensing measurements. Vis/NIR spectral measurements on soil samples can be performed in reflection mode or transmission mode. For that purpose, various spectrophotometers with the same or different resolution are commonly used in the acquisition of soil spectra.

Research conducted in this area has contributed to improve the utility of field collected spectral signatures. Spectrophotometers are being used in commercial agriculture today, but using them to measure soil fertility parameters is relatively new. Consequently, researchers in laboratory settings have found that Vis/NIR reflectance measurements of soil samples correlate to important soil properties (Veris, 2012).

A spectrophotometer is an instrument that consists of a radiant-energy source, a monochromator, sample holder, and detector, used for measurement of radiant flux as a function of wavelength and for measurement of absorption spectra (McGraw-Hill, 2003). According to Vo (2010), with the spectrophotometer, the amount of a known chemical substance (concentration) can also be determined by measuring the intensity of light detected (only substances which are absorbing in the considered wavelength range can be measured).

Soil reflectance is a function of different parameters, such as soil moisture, organic matter, iron oxide and soil texture. Some of these parameters highly correlate with soil spectra, as these are spectrally active in the Vis/NIR range. Therefore, the soil spectral reflectance could be a useful tool for providing a referenced measure of soil fertility parameters. Those data may become in a fundamental component of site-specific-soil management programs, specifically for fertiliser recommendations.

In this chapter the basic and applied concepts of Vis/NIR analyses are illustrated (§ 4.2 to § 4.5). The features of the spectrophotometer used for scanning the soil samples and the optical measurements procedure applied are detailed in section 4.6. The results are presented and discussed in section 4.7. Finally, conclusions are given in section 4.8.

4.2. Interaction of light with matter

Vis/NIR spectroscopy works on the principle of interaction of electromagnetic radiation with matter, which takes several forms. According to Owen (1996) a number of processes can occur when radiation interacts with matter: reflection, scattering, absorbance, fluorescence/phosphorescence (absorption and reemission), and photochemical reaction (absorbance and bond breaking).

In all, reflectance techniques can be divided into two categories: *internal reflection* and *external reflection*. In internal reflection method, interaction of the electromagnetic radiation on the interface between the sample and a medium with a higher refraction index is studied, while external reflectance techniques arise from the radiation reflected from the sample surface. External reflection covers two different types of reflection: *specular* (regular) reflection and *diffuse* reflection. The former usually associated with reflection from smooth, polished surfaces like mirror, and the latter associated with the reflection from rough surfaces (Monsef, 2012).

The incident radiation is reflected (Figure 4-1) back by the outer surface (known as specular reflectance), while another part enters into the inner layers of the sample and eventually leaves the sample at the side of illumination (diffuse reflectance), or the other side of the sample (transmittance), or is absorbed completely (absorption) before it can leave the sample (Figure 4-2).

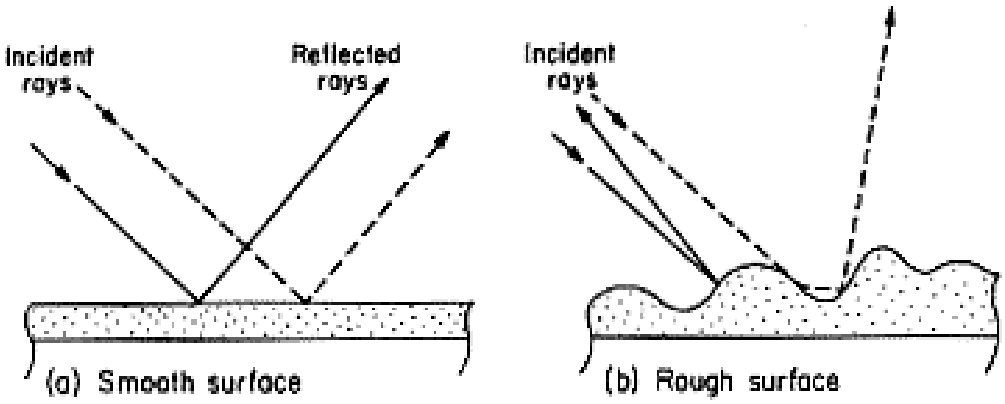


Figure 4-1. Light reflection from a) smooth surface (specular reflection) and b) rough surface (diffuse reflection). In both cases the angle of incidence equals the angle of reflection at the point that the light ray strikes the surface (Source: Optical Society of America, 2008).

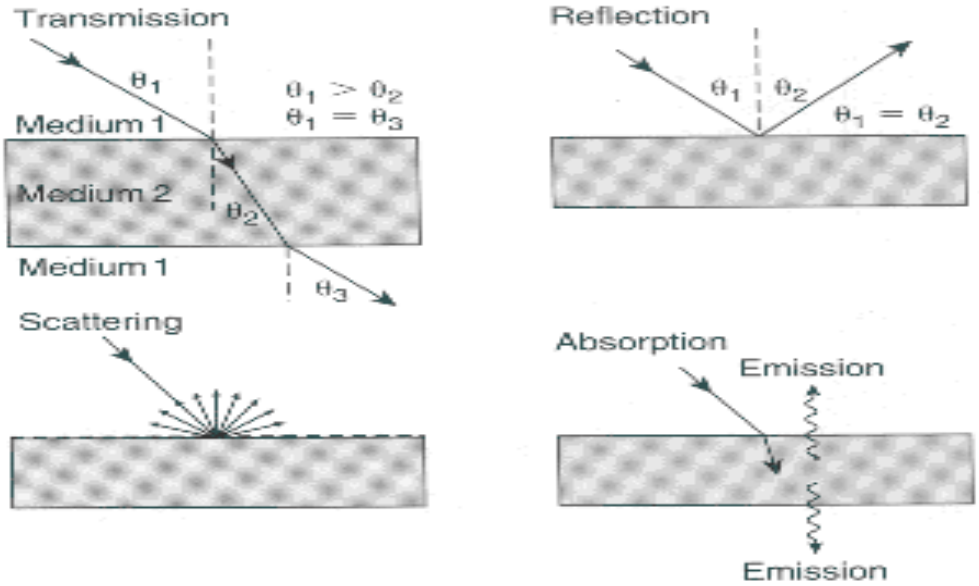


Figure 4-2. Light - matter interactions: Transmission, Reflection, Scattering and Absorption (Source: DeLiberty, 1999).

Interaction of Vis/NIR light with a molecule results in a transformation of the electron configuration (electron transitions or excitation of the bond vibration modes) in the molecular or atomic orbital. The energy of the photons is transferred to the molecule or atom converting it (the electron) to a higher energy state, called an excited state of molecules or atoms. After a short period, the excited molecule is relaxed to its ground state transferring its excess energy to other atoms or molecules.

Guillén et al. (2010) explained that the Vis/NIR spectrum of a molecule is shaped by the absorption at overtones and combination bands of the fundamental molecular absorptions found in the visible and near infrared region. The generally overlapping vibrational bands may appear nonspecific and poorly resolved, making the transmittance or reflectance spectrum difficult to interprete. Therefore, in some cases, it can be very difficult to associate specific characteristics to a determined set of chemical components.

The most common analytical application of spectrophotometry utilizes the relationship between absorbance and concentration. This relationship is described by the Lambert-Bouger-Beer law. This law is a mathematical means of expressing how light is absorbed by matter. The law states that the amount of light emerging from a sample is diminished by three physical phenomena: (i) the amount of absorbing material (concentration C), (ii) the optical path length L, which is the distance the light must travel through the sample, and (iii) the probability that a photon of that particular energy will be absorbed by the sample (Sauer et al., 2008).

$$A = \log\left(\frac{I_0}{I}\right) = kLC \qquad (4.2.1)$$

A is the absorbance, I is the light intensity and I_0 is the incident intensity. A is proportional to the concentration (C) of the absorbing species, L is the pathlength of the light through the solution and k is the absorptivity, a material property (Harris, 2002).

In agreement with the previous statement the spectral signatures of the air-dried soil samples used in this research also were defined by their absorbance, as a function of wavelength in the electromagnetic spectrum. In the case of a sample, which produces diffuse reflectance spectra, the linearity of the Lambert-Bouger-Beer law is corrupted and it can be expressed by the function log (1/R).

According to this law, the concentration of a particular constituent (Birth and Hecht, 1987) is directly proportional to the amount of light absorbed or inversely proportional to the logarithm of the transmitted light. For practical reasons, the diffuse reflectance (R) is converted to absorbance (A) according to the formula:

$$A = \log\left(\frac{1}{R}\right) \qquad (4.2.2)$$

Since different materials absorb light at different frequencies and exhibit different intensity of absorption, there is an interest in determining the amount of various substances in a mixture based on measuring the relative amount of radiant energy absorbed at each frequency. According to Stenberg et al. (2010) the wavelength at which the absorption takes place (i.e., the size of the energy quantum) depends also on the chemical matrix and environmental factors such as neighbouring functional groups and temperature, allowing for the detection of a range of molecules which may contain the same type of bonds.

4.3. Instrumentation

Instrumentation for Vis/NIR spectroscopy is similar to instruments for the UV-visible and mid-IR ranges. There is a source, a detector, and a dispersive element (such as a prism, or, more commonly, a diffraction grating) to allow the intensity at different wavelengths to be recorded. Common incandescent or quartz halogen light bulbs are most often used as broadband sources of near infrared radiation for analytical applications. Light-emitting diodes (LEDs) are also used; they offer greater lifetime and spectral stability and reduced power requirements but limited number of wavebands. The type of detector used depends primarily on the range of wavelengths to be measured. Silicon (Si) - based CCDs (Charge Coupled Devices) are suitable for the shorter end of the NIR range, but are not sufficiently sensitive over most of the range (over 1000 nm). InGaAs (Indium gallium arsenide) and PbS (Lead sulfide) detectors can cover higher wavelength regions than Si detectors, being usual working with both types.

The instrumentation for Vis/NIR spectroscopy can be divided in monochromators, Fourier transform infrared spectrometers and diode arrays. These instruments measure spectral properties, such as absorbance, transmittance and reflectance.

Monochromators are optical devices that disperse light into its spectral wavelengths. These devices form the heart of any optical spectrometric system. More specifically, a monochromator disperses light and isolates one spectral wavelength to be viewed or measured or otherwise used, hence the term monochromator, for monochromatic light. When a wider range of the dispersed wavelengths is used or measured at the same time the same instrument is called a spectrograph. A scanning monochromator, with either a manual or computer controlled motorized turret, scans a selected spectral region one wavelength at a time so that all wavelengths are covered in the end. Scanning monochromators are often used as tunable wavelength light sources or fluorescence excitation sources where monochromatic light is required for illumination. They can also be mated to a single channel detector to measure the spectral content of the input light one wavelength at a time. A combination of a monochromator and single-element detector is a very cost-efficient solution for the measurement of temporally constant spectra.

Fourier transform infrared (FT-IR) spectrometers were developed for commercial use in the 1960's, mainly for advanced research due to the instrument costs and the large computers required for running them. Gradually, technology advancements have reduced the cost and enhanced the capabilities of an FT-IR. Today, FT-IR is the standard for organic compound identification work in laboratories. The working principle of an FT-IR instrument is based on that infrared light emitted from a source is directed into an interferometer, which modulates the light. After the interferometer the light passes through the sample compartment (and also the sample) and is then focused onto the detector. The signal measured by the detector is called the interferogram.

An FT-IR instrument relies upon interference of various frequencies of light to collect a spectrum. The spectrometer consists of a source, beamsplitter, two mirrors, a laser and a detector; the beamsplitter and mirrors are collectively called the interferometer. The modern FT-IR spectrometer has three major advantages over a typical dispersive infrared spectrometer: (i) multiplex advantage; all source wavelengths are measured simultaneously, (ii) throughput advantage; for the same resolution, the energy throughput in an interferometer can be higher (iii) precision advantage; the wavenumber scale of an interferometer is derived from a *helium–neon laser* that acts as an internal reference for each scan.

Diode array spectrophotometers are capable of acquiring complete UV/Visible absorbance spectra in as little as 100 ms. The grating of these instruments is fixed, and rather than moving the grating to acquire spectra, hundreds of detectors are placed at the exit of the monochromator. The detectors are all integrated on a single silicon chip called a photodiode array. The diodes act as capacitors that discharge in proportion to the incident light flux. The capacitance of each diode is converted to a binary word that is input to a computer.

Instruments intended for chemical imaging in the NIR may use a 2D array detector with an acousto-optic tunable filter. Multiple images may be recorded sequentially at different narrow wavelength bands. Many commercial instruments for UV/Vis spectroscopy are capable of recording spectra in the NIR range (to perhaps ~900 nm). In the same way, the range of some mid-IR instruments may extend into the NIR. In these instruments, the detector used for the NIR wavelengths is often the same detector used for the instrument's "main" range of interest.

In Vis/NIR spectroscopy multiple measurements over a certain wavelength range are typically used as they include plenty of information on physical, chemical and biological properties of objects. Commonly, wavelengths ranges are from 350 to 780 nm for Vis, 780-2500 nm for NIR. When Vis/NIR radiation is focused onto a sample, the molecules in the sample will increase their vibration energy by absorbing energy at specific frequencies depending on the molecular geometry, bond strengths and atomic masses. The Vis/NIR lights are thus modified, creating a spectrum or 'signature' of the targeted object with peaks at the absorbing frequencies. The combined contributions from the various soil components can result in a very complex spectrum, difficult to analyse visually, but multivariate calibration models can be built to derive useful qualitative and quantitative relationships or models between the spectral signatures and many soil properties (Yang and Mouazen, 2012).

4.4. Sample presentation

Sample presentation to a NIR instrument is one of the important factors affecting NIR measurements. A variety of sample presentation methods are available to the analytical scientist. These include transmittance (straight and diffuse), reflectance (specular and diffuse), transflection (reflection and transmittance), and interactance (a combination of reflectance and transmittance). Pathlength selection for optimum near infrared

measurements involves the following: for the shortwave near infrared (SW-NIR) region of 800–1100 nm, pathlengths ranging from 5–10 cm are typically used. For longwave near infrared (LW-NIR) or 1100–2500 nm, common pathlengths for hydrocarbons include 0.1–2 cm or 1–20 mm (Workman and Weyer, 2008). Figure 4.3 illustrates these sample presentation mode (Kawano, 2002).

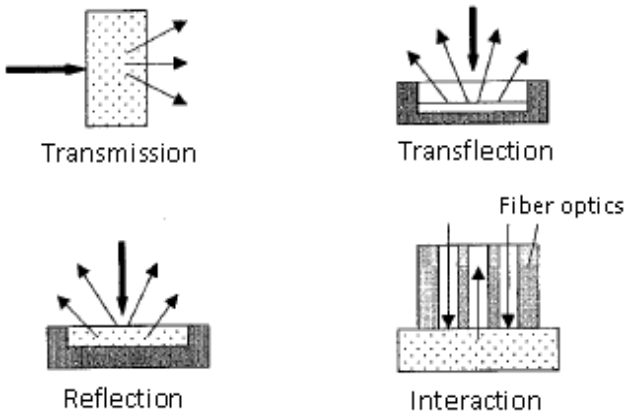


Figure 4-3. Sample presentations of transmission, reflection, transflection and interaction (Source: Kawano, 2002).

Reflectance

Reflection is the process where a fraction of the radiant flux incident on a surface is returned into the same hemisphere whose base is the surface and which contains the incident radiation. In reflection mode the light source and detector are located on the same side of the instrument (Figure 4-3). The reflection can be specular (in the mirror direction), diffuse (scattered into the entire hemisphere), or a combination of both.

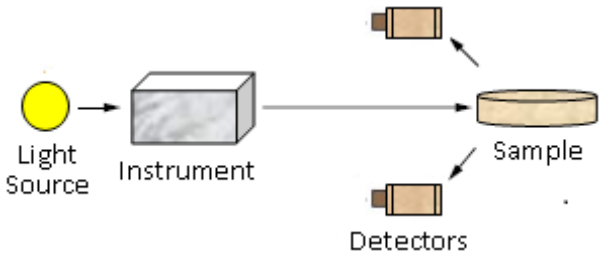


Figure 4-4. Basic instrument configuration for reflectance measurement.

In line with Karavanova (2001) each kind of objects has its own specific reflectance characteristics: type of spectral curves and different values of the spectral coefficients in the different bands. These characteristics are determined by the physical-chemical properties of the objects.

Transmitance

In the case of transmission, incident light illuminates one side of the sample and the transmitted light may be detected from the other side. This presentation is widely used for liquids (Kawano, 2002). In transmission mode the light source and detector are on opposite sides of the instrument (Figure 4-4). This is critical for quantitation of absorbances and requires an adapter for transparencies or negatives to be mounted or built-in on the scanner.

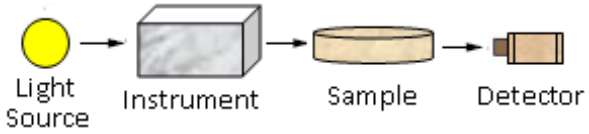


Figure 4-5. Basic instrument configuration for transmittance measurement.

In transmittance measurements the entire pathlength of a sample is integrated into the spectral measurement, thereby reducing errors due to nonhomogeneity of samples. Transmittance techniques are most useful for measuring transparent liquids and large particles. For fine particles, the front surface scatter brings about a loss of energy transmitted through a sample with the net effect being a decrease in the signal-to-noise ratio. In transmittance, higher frequency energy is most commonly used due to its greater depth of penetration into the sample. The higher frequency energy (800 to 1400 nm) is more susceptible to front surface scattering than lower frequency energy. Transmittance measurements must therefore be optimized, taking into consideration the relationships between the frequencies used for measurement, front surface scatter, and the pathlength of the sample. In transmittance measurements, particle size can be small enough to begin to scatter most of the energy striking the sample. If the particle size is sufficiently small, the instrument will not transmit enough energy through the sample for the detectors to record a signal. For this, the ideal research instrument would have both transmittance and reflectance capabilities.

Transflectance and interactance

Transflection was originally developed by Technicon (American scientific instruments company) for the InfraAlyzer and combines transmission and reflection. Incident light is transmitted through the sample and then scattered back from a reflector, which is made of ceramic or aluminium to be compatible with the diffuse reflection characteristics of the instrument.

In the case of interactance, an interactance probe having a concentric outer ring of illuminator and an inner portion of receptor is usually used. The end of the probe should be in contact with the surface of the sample. Therefore only the light transmitted through the sample can be detected.

4.5. Interaction of light with soil

In general the spectral reflectance of soil is determined for the larger part by six variables: moisture content, organic matter content, particle size distribution and iron oxide content, soil mineralogy and soil structure. Of these variables, moisture content is the most important due to its dynamic nature and large overall impact on soil reflectance. The spectral reflectance curves could be also influenced by the soil colour. Colour is a spectral property, which itself is caused by the chemical composition (absorption) and microstructure properties (scattering).

Latz et al. (1984) reported that two important properties affecting the spectral reflectance of soils are the organic matter and iron oxide contents. High organic matter contents tend to cause the reflectance to be low and give a concave character to the spectral response. High iron oxide content along with the reduction in organic matter tends to cause an increase in the overall spectral response. However, the presence of iron oxides causes a leveling off of the spectral response. Intermediate contents of organic matter and iron oxide affect the spectral response relative to their proportions within the soil sample.

Also, Karavanova (2001) explained that for soils the most important properties that influence the level of reflectance are: humus content, salinity, moisture, structure of the arable horizon and content of carbonates, ferric hydroxides and gypsum ($CaSO_4 \cdot 2H_2O$).

Bogrekci and Lee (2005b) stated that reflectance from a soil sample varies with regard to soil particle size in UV-Vis/NIR region of the electromagnetic spectrum. This photo-physical effect of the soil sample is very useful in determining the particle size distribution. However, the same feature is not desirable in sensing P concentration of a soil sample due to the fact that this photo-physical property introduces variation of the reflectance intensity at different wavelengths. Therefore, the effect of the particle size on the reflectance spectra of a soil sample should be taken into account when the chemical properties of the soil sample are derived from these spectra.

Reeves et al. (2005) explained that studies on the occurrence and effects of specular reflection in mid infrared spectra of soils have shown that distortions due to specular reflection occur for both organic (humic acid) and non-organic fractions (carbonates, silica, and ashed fraction of soil). Diffuse reflection occurs when light shines onto a powder sample like soil; it is reflected in all directions, as shown in Figure 4-6.

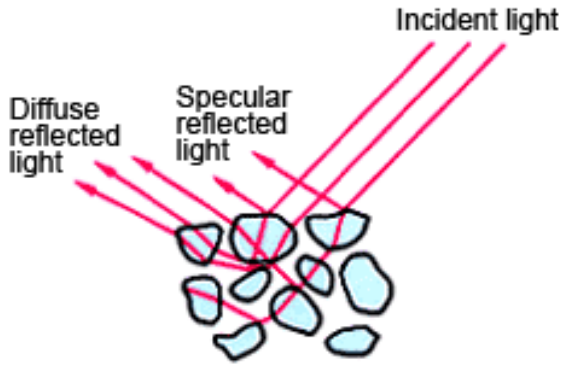


Figure 4-6. Schematic diagram of light scattering from a powder sample (Source: Shimadzu, 2012).

The specular (or mirror) reflection causes the shiny appearance of a flat surface, whereas external diffuse reflectance is induced by rough surfaces. Specular reflection and external diffuse reflection do not contain information on the chemical composition of the sample, while internal diffuse reflectance does provide this information as the light has interacted with the matter and thus had the chance to be absorbed.

Some of the light undergoes specular reflection at the powdered surface. Due to the variety of powder shapes, it is reflected in many directions, unlike light reflected from a mirror. The remainder of the light is refracted as it enters the powder, where it is scattered due to internal reflection. Some of this scattered light is emitted back into the air (Shimadzu, 2012).

As the diffuse reflected light is reflected or passes through the powder, it becomes weaker if absorption by the powder occurs. This results in a diffuse reflected spectrum, similar to the transmission spectrum. However, in regions where the powder exhibits strong absorption, most of the diffuse reflected light in long light paths is absorbed, such that only the diffuse reflected light from short light paths is emitted back into the air. Conversely, in bands of weak absorption, some light is not absorbed even in long light paths and this diffuse reflected light is emitted to the air. When there is high absorption,

one needs a very thin sample to measure some transmittance, while one would still measure diffuse reflectance if the sample is sufficiently scattering.

In such a diffuse reflectance spectrum, the absorbed wave number positions are the same as the transmission spectrum (Shimadzu, 2012).

Experimental results obtained by Leblon (1997) showed that the soil reflectance was increased with decreasing particle size for a given type of soil. In other research carried by Sun et al. (2009) noted that the curves of the soil spectral reflectance decreased, they considered the possible influences of soil texture and colour. Also, in that study, the valleys of the spectral reflectance appeared at 1420, 1910 and 2210 nm. Discussing the spectral reflectance of different types of soil texture, it was easy to find that the reflectance of fine grained soil was higher than of the rough grained soil.

In agreement with Ben-Dor et al. (1999) soil OM also could influence soil spectra in the visible range. Mutuo et al. (2006) found that OM fractions exhibited strong absorption in the region 500 – 850 nm. This result was attributed to the reported by Schubert (1965) who related to the edge of a large absorption feature caused by lignin centred at 280 nm.

4.6. Materials and methods

4.6.1. Equipment

The setup for acquiring Vis/NIR diffuse reflectance spectra of soil samples in the laboratory is illustrated in Figure 4-7. It consists of a diode array spectrophotometer (CORONA PLUS REMOTE Vis/NIR SB, Zeiss, Jena, Germany) and an OMK 500-H measuring head connected to it with an optical fibre-bundle.

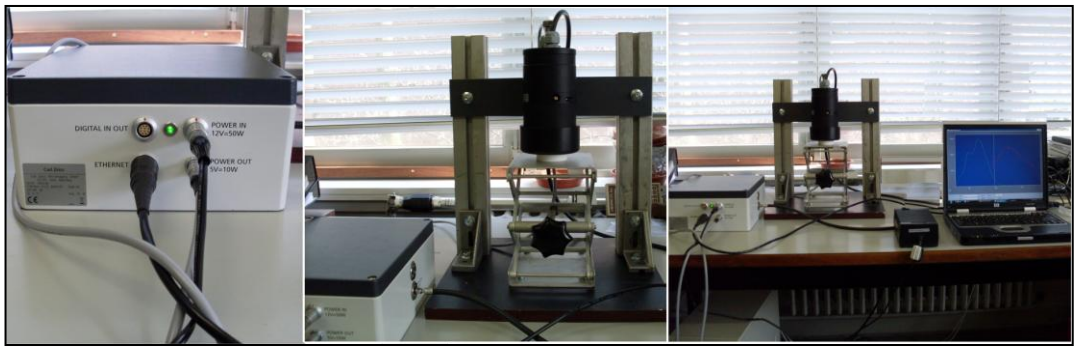


Figure 4-7. Spectrophotometer CORONA PLUS REMOTE VIS-NIR SB (left), OMK 500-H measuring head (middle), and entire setup including the laptop computer (right).

The OMK 500-H measuring head has been designed for measuring dense scattering, relatively heterogeneous samples in diffuse reflectance mode. It is especially suitable for on-line measurement of the diffuse reflectance of samples with continuously varying distance between the measuring head and sample surface.

The specifications of the OMK 500 H NIR measuring head are summarised in Table 4-1. It contains a 5 V, 10 W halogen lamp powered by a power supply through an electric cable. An optical system provides sample illumination at 0° (normal to the sample) in a quasi-parallel beam. In the colour measuring head, 15 individual optical fibres are uniformly arranged in a ring for sample observation at 45°.

The fibres are bundled to a light guide that is connected to the spectrometer. Additionally, the light guide connection for the measuring channel is located on the front of the instrument. The computer (PC) is connected to the spectrophotometer via an Ethernet interface. The minimal configuration is determined by the software used. No operations on the device are required, as control is handled via the software, on the PC.

Various software packages are available for this measuring system. In this research the Aspect –*plus* software supplied by Carl Zeiss Jena GmbH was used.

Table 4-1

Technical data of the OMK 500-H-NIR measuring head.

OMK 500 measuring heads **Designation**	**Remarks**
	OMK 500-H NIR
Spectral range	400 nm ... 2 200 nm
Measuring time per sample	< 2 s
Measurement range	950 ... 2 050
100 % Noise	Δ R < 0.04 % R
Sensitivity at varying distance	X ± 1 mm dR < 0.2 %
Integration time for full scale deflection	< 100 ms
Light source	halogen lamp 5 V, 10 W, stabilized
Lifetime of lamp	approx. 3 000 hours
Fibre connector	2x Zeiss-connector
Viewing geometry	15 x 24° circular under 45°
Illuminating angle	0°
Sampling spot size	approx. 20 mm
Length of optical fibre	1 m ... 10 m
Weight	1.3 kg
Dimensions	100 mm x 100 mm x 197 mm (W x D x H)
Enclosure protection	IP 55

4.6.2. Spectral data acquisition

The Vis/NIR reflectance spectra were acquired in the laboratory for all the mixture soil samples collected from different fields (landscape level) and for the different subsamples taken at the selected Cambisol and Vertisol field (field level). All spectra were obtained using the same instrument settings: average 10 scans per spectrum, measure in reflection mode, for wavelength range from 399 to 1697 nm. The integration times for the Vis (Si: 400 – 1000 nm) and the NIR (InGaAs: 900 – 1700 nm) ranges were set to 54 ms and 43 ms, respectively. The spectral resolution in Vis was 3.3 nm and in NIR was 10 nm.

A small amount of the soil sample (about 30 g) was placed in a petri dish of 10 mm depth and 35 mm diameter. The soil in the petri dish was first compacted and then carefully levelled in order to obtain a smooth surface to reduce variation due to the packing of the soil. Each petri-dish with 30 g of soil was placed under the soil sensor at the focal point. Then three reflectance spectra were taken over the central area of the petri dish rotating the sample approximately 120° between each spectral acquisition.

The three spectra of each soil sample were averaged to obtain one average spectrum per sample. Each spectrum was saved as an individual file (.csv), and then these files were assembled into a single matrix to be imported into MATLAB.

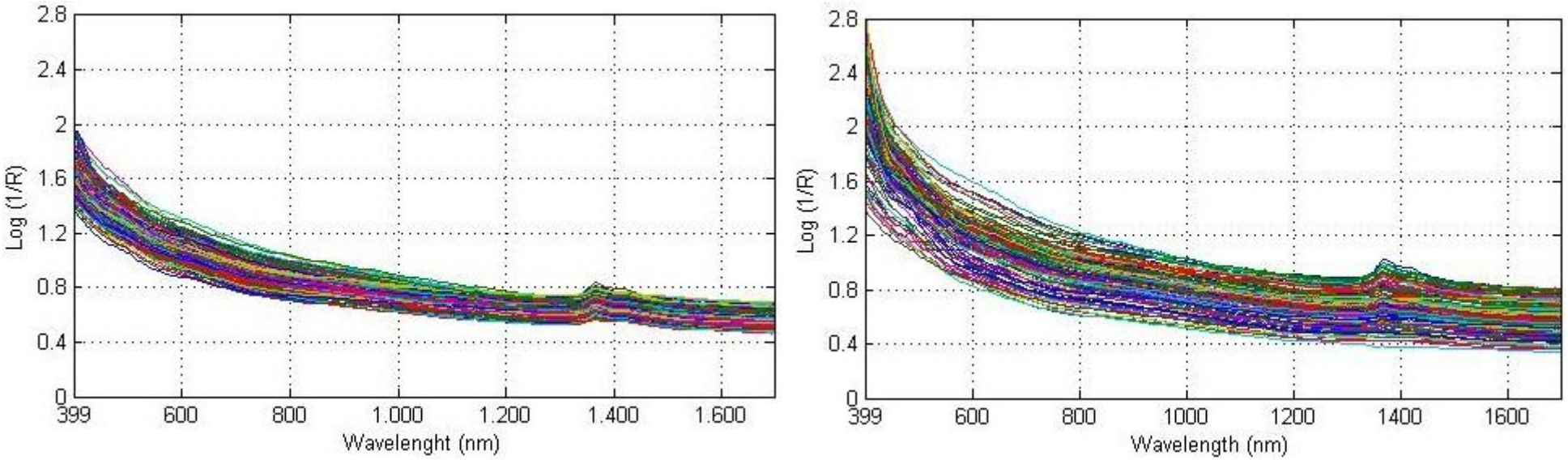
4.7. Results and discussion

4.7.1. Spectral absorbance at landscape level

The absorbance curves are a graphical representation of the spectral absorbance of the air-dried soil samples, as a function of the studied wavelength range comprised in the Vis/NIR region. The absorbance spectra of these two soil types can clearly be distinguished, which suggests that the soil type could be derived from the spectra.

The absorbance spectra for these two soil groups are illustrated in Figure 4-8 A (Cambisol) and 4-8 B (Vertisol). The light absorbance by the soil varied with wavelength in each soil group. In addition, the pattern of soil signatures for each soil group at landscape level was very similar. According with it the absorbance decreased as wavelength increased.

However, the spectra of both types of soil samples have the same general characteristics. In the visible region (399 – 780 nm) the absorbance was higher than in the NIR region (780 – 1697 nm). Also, the maximum values of absorbance at the wavelength of 399 nm for both groups. Magnitudes of absorbance of these soil signatures for each group were different. Around the wavelength of 1450 nm in both soil groups the absorbance values increase and consequently decrease again around the same wavelength. For Cambisol the absorbance values were generally lower than Vertisol. On the other hand the absorption features around 1450 nm were highlighted in both soil groups.



A **B**

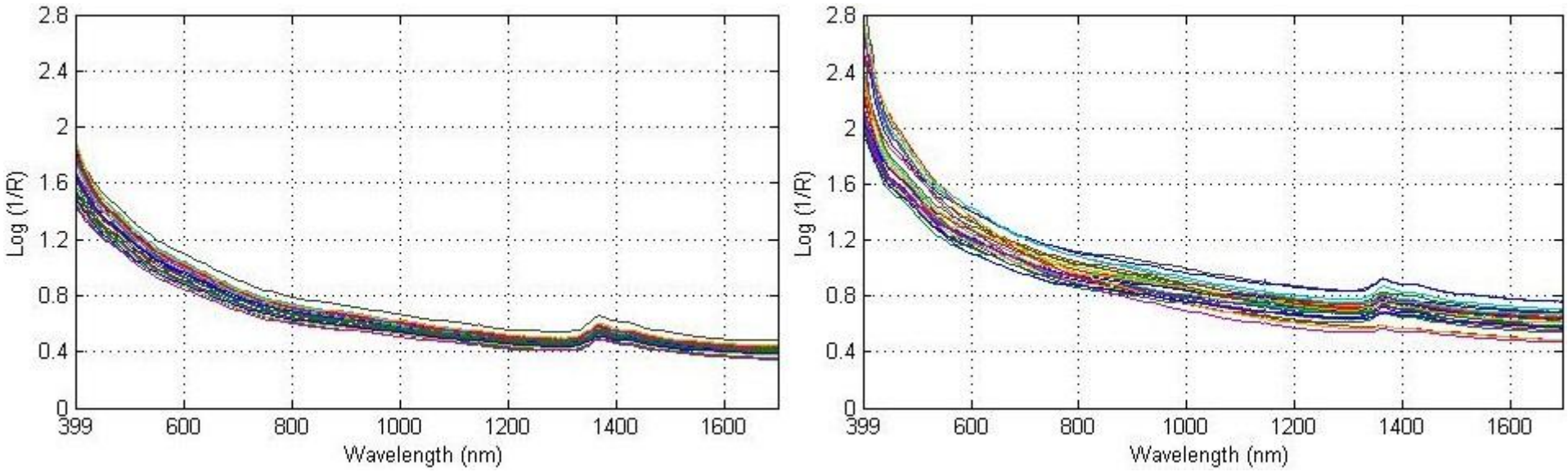
Figure 4-8. Absorbance spectra for the soil samples collected at landscape level
A- Cambisol B- Vertisol.

4.7.2. Spectral absorbance within a field

In general, within a field the absorbance of the air-dried soil samples decreased as wavelength increased and the maximum values were observed at the wavelength of 399 nm, corresponding to the visible region. This pattern was very similar for Cambisol and Vertisol (Figures 4-9 A and 4-9 B respectively). In general, for the first group the absorbance values were lower. Also, in the NIR region the spectral absorbance for Cambisol and Vertisol field showed a strong peak around the wavelength of 1400 nm and then, decrease toward 1697 nm.

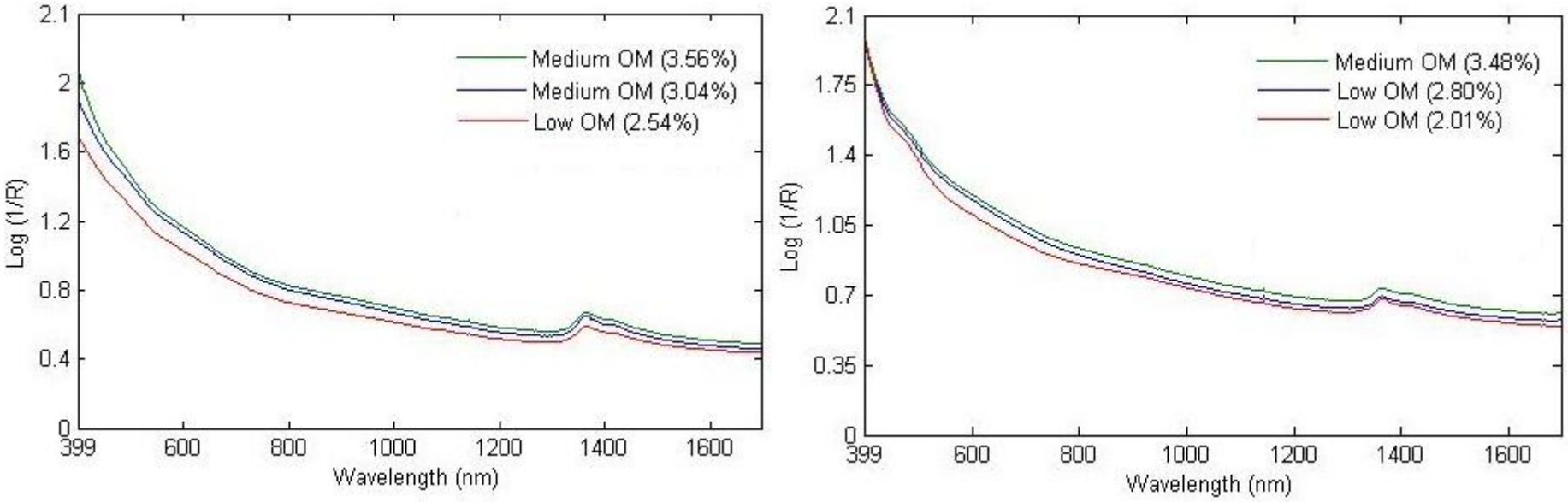
A visual comparison of the spectral shape and the variation between a field and at landscape level within the same soil group showed some similitude. This similitude is related to the absorbance values, which also decreased from the Vis to the NIR region. On the other hand a similar peak was observed around the 1400 nm, which means that the absorbance curves peak in water absorption bands. These characteristics bring a good match of the spectral signatures within a field and at landscape level. Even dry, clayey soil dips in water absorption bands around the 1400 nm as the clayey soil has water molecules in its chemical composition and not all water molecules are removed by the process of air drying. This last statement is in line with Choudhury et al. (2009).

In both types of soils an increasing pattern of soil absorbance was observed as OM% increases. For Cambisol type this pattern was observed already at 399 nm, while for Vertisol type it was more clearly observed around 580 nm (Figures 4-10 A and 4-10 B respectively). The relationship between increasing OM content and decreasing reflectance exits primarily for soils with more than 1.5% OM content, above this show a decrease in reflectance with increase in OM. On the other hand when soils have less than 1.5%, iron oxide content of the soil exerts significant effect on the level of reflectance (Montgomery, 1976). Thus, Irons et al. (1989) considered that an increase in the OM content of a soil generally causes a decrease of reflectance over the entire spectrum. A high OM content and hence, a strong decrease of overall reflectance, might even mask other absorption features in the soil spectra.



A **B**

Figure 4-9. Absorbance spectra for the soil samples collected within a field
A- Cambisol B- Vertisol.



A B

Figure 4-10. Soil absorbance spectra with different OM contents within a field
A- Cambisol B- Vertisol.

4.8. Conclusions

First, the interaction of light with matter and the possibility to extract information on the chemical composition of that matter from the changes in the light spectrum have been discussed. Secondly, the equipment and sample presentation strategies for acquiring these spectra have been described. Thus, a variety of sample presentation methods such as transmittance, reflectance, transflection and interactance are available to the analytical scientist.

The instrumentation used for Vis/NIR spectroscopy measure spectral properties, such as absorbance, transmittance and reflectance. These instruments are divided in monochromators, Fourier transform infrared spectrometers and diode arrays. These measurements over a wavelength range include important information on physical, chemical and biological properties of soils. In general the spectral reflectance of soil is determined for different parameters such as moisture content, organic matter content, particle size distribution and iron oxide content, soil mineralogy and soil structure. Therefore the combined contributions from these soil components result in a complex spectrum. It is important to take into account that sample presentation to a NIR instrument is one of the important factors which affect NIR measurements. Finally, the setup used in this research has been described and the soil spectra obtained with this have been discussed.

From the acquired reflectance spectra distinct differences were observed between the different soil types (Vertisol and Cambisol) and between soil samples from different fields within one soil group. Even within a field clear variation was observed between the acquired reflectance spectra. The reflectance values observed in the Vis region were lower than those in the NIR region.

Chapter 5

Vis/NIR measurement of different P levels added to the same soil

5.1. Introduction

Reflectance measurements of soils in the Vis/NIR range to investigate the soil properties have been conducted by many researchers (Ben-Dor et al., 1999; Bogrekci and Lee, 2005a; Bogrekci and Lee, 2005b; Bogrekci and Lee, 2005c; Aïchi et al., 2009; Cécillon et al., 2009). Shepherd and Walsh (2002) proposed a spectral library approach, whereby the variability of soils in a study area is thoroughly sampled and spectrally characterized. Soil properties or attributes of soil quality are measured on only a selection of soils, designed to sample the variation in the spectral library, and then calibrated to soil Vis/NIR reflectance. This approach can be extended to provide spectral indicators of soil quality.

One of the advantages of Vis/NIR technology is not only to assess chemical structures through the analysis of the molecular bonds in the Vis/NIR spectrum, but also to build a characteristic spectrum that represents the ‘‘finger print’’ of the sample (Cozzolino and Murray, 2004). In line with Bogrekci and Lee (2005c) reflected light from soil carries physical and chemical information about the material with which it has interacted. Most spectroscopy studies on soils have focused on the visible and near infrared regions of the electromagnetic spectrum due to many reasons such as cost, availability, and sensitivity. However, in the study developed by these authors, the ultraviolet (UV) region was added to explore more possibilities of developing better P prediction models.

The electromagnetic spectrum of a soil sample holds information about soil, water content, nutrients, and particle size. Therefore, the determination of soil spectral signatures improves the prediction accuracy of calibration models for the determination of soil P concentrations (Bogrekci and Lee, 2005a).

This experimental study aimed to investigate the effect of different levels of total phosphorus (P) on the spectra of five types of soil.

5.2. Materials and methods

For this study, a designed experiment split in two parts was carried out. The first part was designed to add different P levels to several soil samples. The topsoil samples (0-15 cm depth) used for this experiment were collected from fallow land in agricultural areas, at five different sites, three from Kenya and two from Madagascar (Table 5-1 and Figure 5-1). The soils were air-dried and sieved over a 4 mm sieve. All soils were characterized by a low amount of available P (AEM[6] P ranges between 1 and 8 mg P kg^{-1}), but had contrasting P sorption indices (Six et al., 2012).

Table 5-1

General characteristics of the five soils studied (Data source: Six et al., 2013).

Soil	Soil type[a]	Sand[b]	Silt[b]	Clay[b]	pH[c]	PSI[d] (mg kg^{-1})
		(%)				
Teso (Kenya)	Cambisol	82	12	6	4.6	15
Betafo (Madagascar)	Andosol	54	28	18	4.3	1070
Sega (Kenya)	Ferralsol	23	12	65	4.2	140
Ivory (Madagascar)	Ferralsol	53	4	43	4.4	120
Kuinet (Kenya)	Ferralsol	16	6	78	4.2	220

[a] major soil grouping (FAO, 1990); [b] particle size analyses by pipette method (Day, 1965); [c] pH (1:5) determined in 0.01 *M* $CaCl_2$; [d] P sorption index (PSI) defined as the amount of P sorbed on the solid phase at a soil solution concentration of 0.2 mg P L^{-1}.

Teso Betafo Sega Ivory Kuinet

Figure 5-1. Soil types used for the designed experiment.

For adding different P levels to the soil samples (Table 5-2), first 750 g of each soil type was taken and split into 5 samples (150 g per sample). Then, the ideal mixing moisture content (MC) of each soil type was tested by trial and error in 50 g of soil placed in petri dishes. In this case, soils needed to be wet, but not too wet to allow good mixing. According to that, the quantities of H_2O for 50 g of soil was set in 3.4 ml (Teso), 12.4 ml

[6] AEM- Anion Exchange Membrane

(Betafo), 8.2 ml (Sega), 6.8 ml (Kuinet) and 5.3 ml (Ivory). Thus, based on the rule of three the quantities of H_2O for 150 g of soil were set in 10 ml (Teso), 37 ml (Betafo), 25 ml (Sega), 16 ml (Ivory) and 20 ml (Kuinet). For adding different P levels it was necessary to prepare a stock solution of KH_2PO_4 (2.5 g P L^{-1} = 2500 mg P L^{-1}). For obtaining the g of KH_2PO_4 needed it was calculated the ratio between the molecular weight (MW) of P (31 g mol^{-1}) and KH_2PO_4 (136 g mol^{-1}). With the 2500 mg P L^{-1} considered to prepare the stock solution and this ratio value (0.23) 10.97 g of KH_2PO_4 was necessary to prepare the stock solution. The amount of mg P for 150 g (0.15 kg) of soil was determined. Then, this amount of mg P for 0.15 kg^{-1} of soil needed was converted to ml from the stock solution and the remaining moisture was added as H_2O based on the MC of each soil type. The soil samples were incubated for two weeks, and then were oven-dried during 48 h.

Table 5-2

Equivalence between P levels and ml from the stock solution and H_2O.

P level (mg P kg^{-1})	P level (mg P 0.15 kg^{-1})	P stock (ml)	H_2O (ml)				
			Cambisol	Andosol	Ferralsol (Sega)	Ferralsol (Ivory)	Ferralsol (Kuinet)
0	0	0	10	37	25	16	20
20	3	1.2	9	36	23	15	19
40	6	2.4	8	35	22	14	18
80	12	4.8	5	32	20	11	16
160	24	9.6	1	28	15	6	11

The second part was designed to obtain the spectral reflectance curves from these soil samples with different P levels added. The five samples (150 g) with different P levels were split into five replicates corresponding to five subsamples of about 30 g each. These subsamples were placed into petri-dishes of approximately 10 mm depth and 35 mm diameter. The soil in the petri dish was first pressed and then carefully levelled in order to obtain a smooth surface to reduce variation due to the packing of the soil. The setup for acquiring Vis/NIR diffuse reflectance spectra of soil subsamples in the laboratory is illustrated in Figure 5-2. It consists of a diode array spectrophotometer (ANALYTICAL SPECTRAL DEVICE INC., USA) and a sensor connected to it with an optical fibre-bundle.

Figure 5-2. Spectrophotometer ANALYTICAL SPECTRAL DEVICE, INC (left), sensor scanning a subsample (middle), and entire setup including the laptop computer (right).

The Vis/NIR spectra for all subsamples were acquired using the same instrument settings: average 10 scans per spectrum, measure in reflection mode, wavelength range from 350 to 2500 nm. For the Vis (Si) and the NIR (InGaAs) ranges all the spectra were acquired with less than 100 ms integration time per spectrum. Each petri-dish with the 30 g of soil sample was placed under the soil sensor at the focal point. Then for each subsample three reflectance spectra were taken over the central area of the petri dish at the same position without rotating the sample. The three spectra were saved as individual files (.csv), and then those files were assembled into a single matrix for MATLAB 7.9 (R2009b, The Mathworks, Nattick, MT). In total 75 spectra were recorded in each soil type, which corresponded to 3 spectra per subsample and 15 spectra per sample for each one of the P levels.

Currently, it is common practice to assess the predictive ability of multivariate models by comparing predictions with reference values for a test set. From the squared deviations, a root mean squared error of prediction (RMSEP) is calculated as:

$$RMSEP = \sqrt{\frac{1}{N}\sum(\hat{y}_{\mathrm{i}} - \mathrm{y}_{\mathrm{i,ref}})^2} \qquad (5.2.1)$$

Where N denotes the size of the test set, and $\hat{y}$ and $y_{i,ref}$ are the prediction and reference value for sample i, respectively.

The RMSEP was determined by using Partial Least Squares (PLS) regression. These calculations were performed in MATLAB 7.9. For every RMSEP 60 soil spectra were used in the PLS calibration and the remaining 15 spectra for validation. This process was repeated five times (Figure 5-3).

The pre-processing method (SNV, Mean Centre) was applied. The Venetian blinds cross-validation strategy was used on the calibration set to decide on the model complexity (number of latent variables) and to evaluate this prediction error.

	(1)		(2)		(3)		(4)		(5)	
	X	**Y**	**X**	**Y**	**X**	**Y**	**X**	**Y**	**X**	**Y**
Test set	.....	0	.....	20	.....	40	.....	80	.....	160
Calibration set	.....	20	.....	0	.....	0	.....	0	.....	0
	.....	40	.....	40	.....	20	.....	20	.....	20
	.....	80	.....	80	.....	80	.....	40	.....	40
	.....	160	.....	160	.....	160	.....	160	.....	80

Figure 5-3. Calibration and validation set for obtaining the RMSEP

5.3. Results and discussion

Figures 5-4 to 5-6 show the average reflectance spectra in the Vis/NIR range for the samples of five soil types with different levels of added phosphate. For soil type Cambisol (Figure 5-4 a) the P level 0 mg P kg^{-1} showed the highest reflectance values in the Vis and NIR region. These values were increasing from the Vis to the NIR region. The highest reflectance was around 2070 nm for all the P levels studied. Around this wavelength of 2070 nm the reflectance values begin to decrease until 2500 nm. The reflectance values for all P levels sharply decreased around 1400; 1900 and 2250 nm, which corresponded to water absorption bands at these wavelengths.

The Andosol also showed the highest reflectance values at P level 0 mg P kg^{-1}, however with a little difference with respect to the other P levels (Figure 5-4 b). Also, these values were increasing from the Vis to the NIR region. The highest reflectance was around 1850 nm for all the P levels studied, and then begins to decrease until 2500 nm. In this soil type the spectra dip for water absorption bands occurs near to the same wavelengths observed in Cambisol.

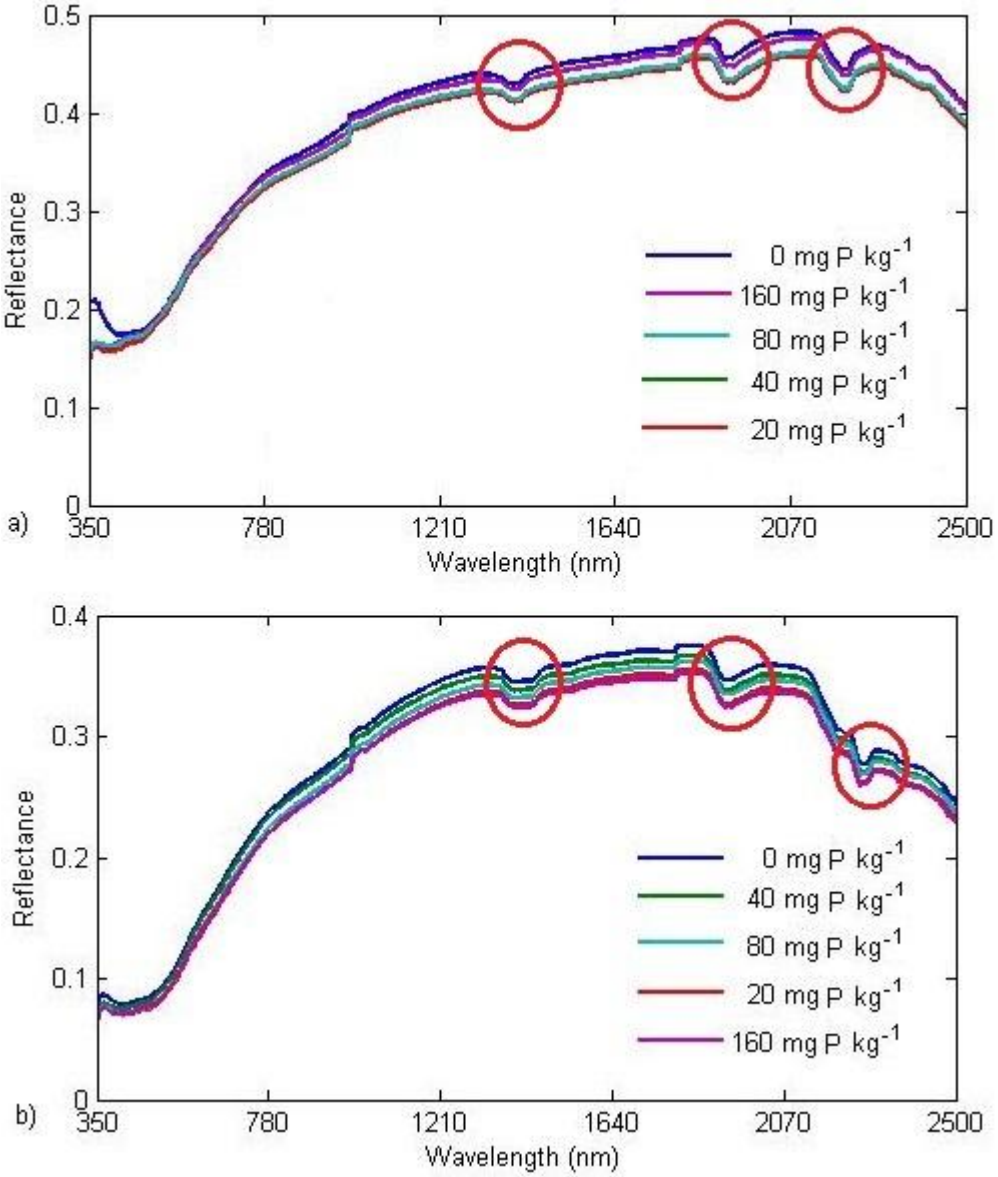


Figure 5-4. Vis/NIR Spectral variation of different P levels
a) Cambisol; b) Andosol.

O- Indicate the water absorption bands

In the three Ferralsols studied (Sega, Ivory and Kuinet) the highest reflectance was not observed at the P level 0 mg P kg^{-1} but in Ferralsols Sega and Kuinet, it corresponded to the added P content of 80 mg P kg^{-1}. In Ferralsol Ivory, the higher reflectance was observed at P level of 40 mg P kg^{-1}, while the lower corresponded to the P level of 0 mg P kg^{-1} (Figure 5-5).

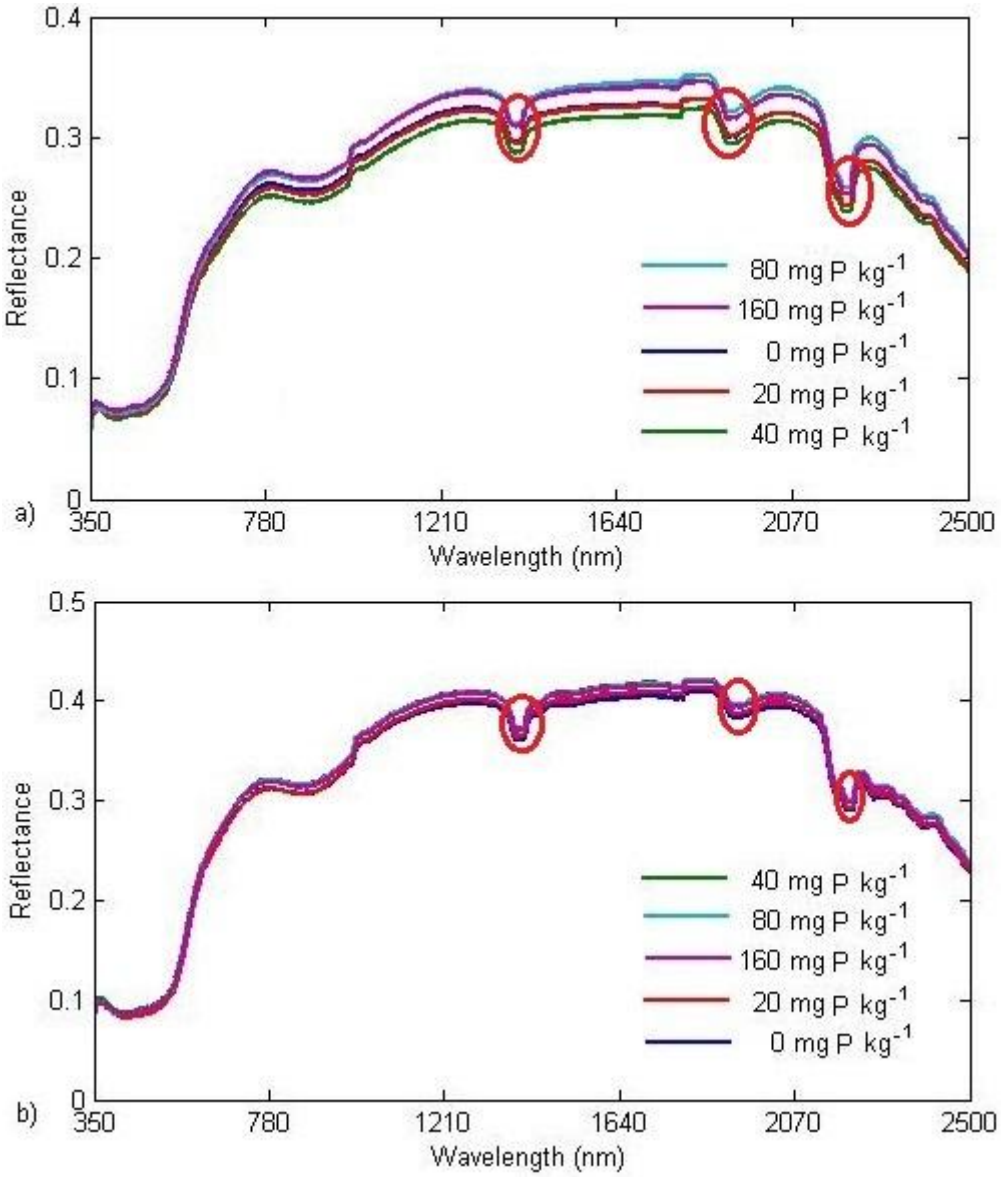


Figure 5-5. Vis/NIR Spectral variation of different P levels a) Ferralsol (Sega); b) Ferralsol (Ivory).

O- Indicate the water absorption bands

In Ferralsol Kuinet, the higher reflectance was observed at the higher P level of 160 mg P kg^{-1}, while the lower corresponded to the P level of 20 mg P kg^{-1}. These values were increasing from the Vis to the NIR regions. However, it showed almost the same reflectance from the wavelengths around 1210 nm to the 1850 nm (Figure 5-6).

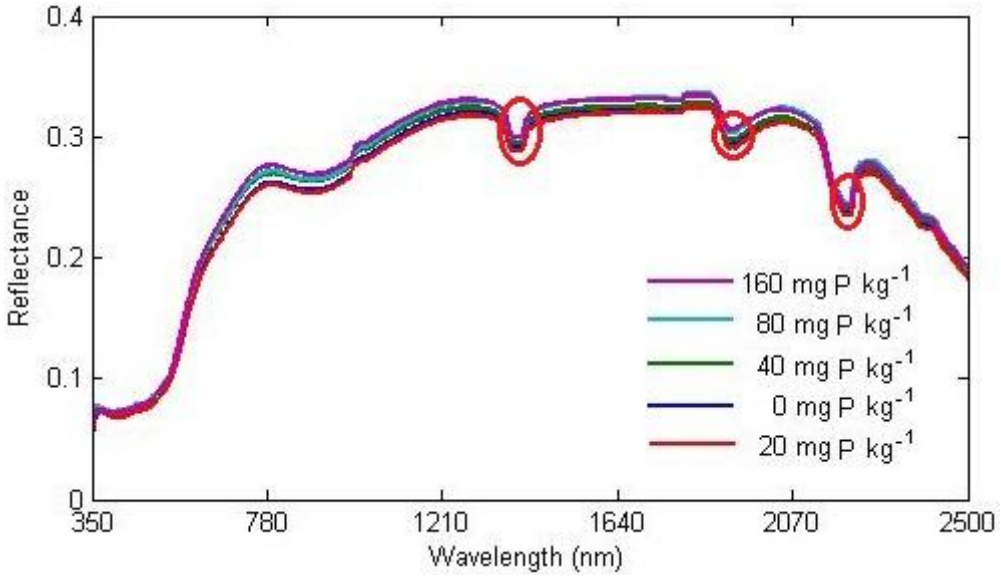


Figure 5-6. Vis/NIR Spectral variation of different P levels Ferralsol (Kuinet).

O- Indicate the water absorption bands

A similar spectral dip at the water absorption bands as for Cambisol and Andosol was observed in all the Ferralsols (Figures 5-5 a; 5-5 b and 5-6). Although the soil was oven dried, it was still possible to observe a water signature in the spectra. In agreement with Bogrecki and Lee (2005a) this difference in reflectance for each soil type could originate from the difference in particle size. According to these authors the measured reflectance spectrum is the result of a combination of soil properties that include particle size, water, organic matter, and nutrients. When nutrient content, organic matter, water and particle size change, the reflectance spectra of the same soil will change.

The Vis/NIR predictions of different P levels added to the same soil are shown in Figures 5-7 to 5-21. For the Cambisol the RMSEP values decreased from the P level 0 mg P kg^{-1} to 20 mg P kg^{-1}. This result indicated that the power of the predictive model was improved by adding this P quantity (20 mg P kg^{-1}) to the soil samples. The P level 0 mg P kg^{-1} tends to be under predicted while the level of 20 mg P kg^{-1} is over predicted (Figure 5-7).

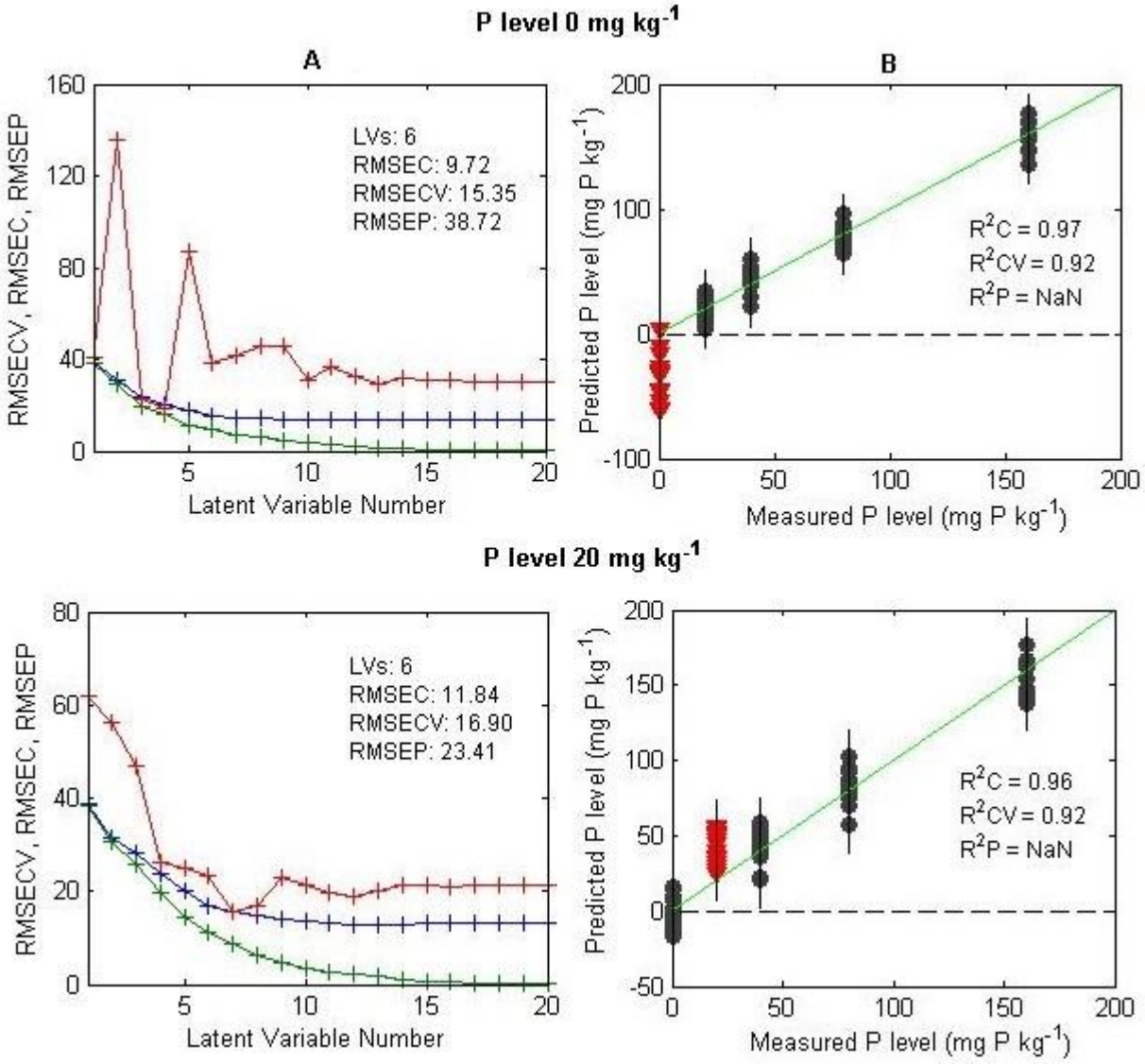


Figure 5-7. Vis/NIR prediction of different P levels (0 and 20 mg P kg^{-1}) added to the Cambisol samples A- SIMPLS Variance captured and Statistics B- PLS regression model.

LVs- latent variables; RMSE- root mean square error; R^2-Coefficient of determination
C- Calibration; CV- Cross Validation; P- Prediction
NaN- Not a Number

Although the RMSEP values gradually increased from 40 mg P kg^{-1} to 80 mg P kg^{-1}, these values showed some similarities with respect to the observed in 20 mg P kg^{-1}. Also, are comparable with their respective RMSECV as in 20 mg P kg^{-1} (Figure 5-8).

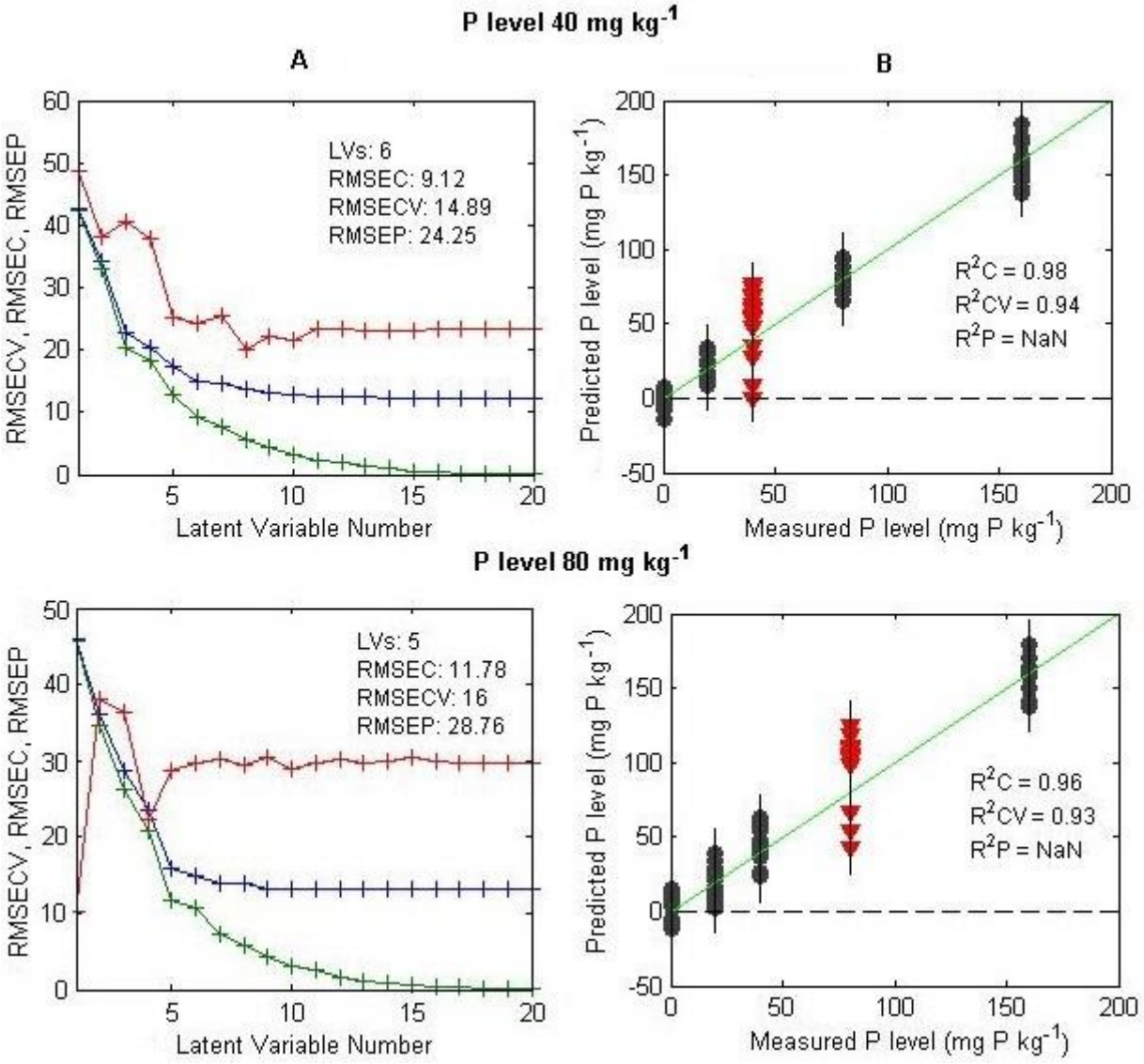


Figure 5-8. Vis/NIR prediction of different P levels (40 and 80 mg P kg^{-1}) added to the Cambisol samples A- SIMPLS Variance captured and Statistics B- PLS regression model.

LVs- latent variables; RMSE- root mean square error; R^2-Coefficient of determination
C- Calibration; CV- Cross Validation; P- Prediction
NaN- Not a Number

The RMSEP at 160 mg P kg^{-1} is completely extrapolated showing a high difference with respect to the RMSECV. This added P level was completely under predicted (Figure 5-9).

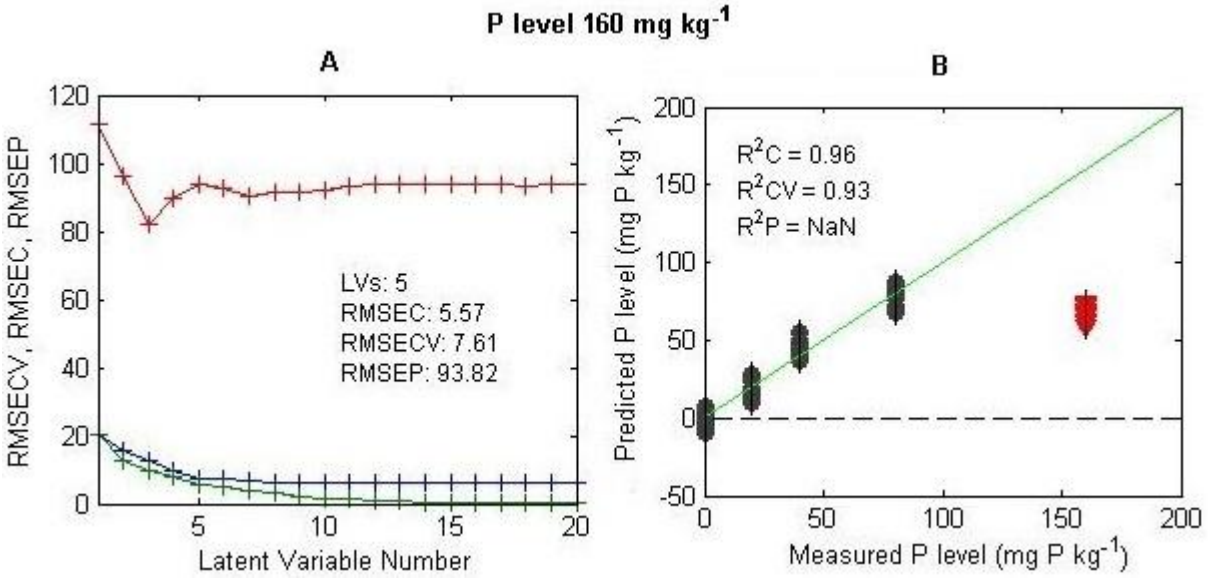


Figure 5-9. Vis/NIR prediction of different P levels (160 mg P kg^{-1}) added to the Cambisol samples A- SIMPLS Variance captured and Statistics B- PLS regression model.

LVs- latent variables; RMSE- root mean square error; R^2-Coefficient of determination
C- Calibration; CV- Cross Validation; P- Prediction
NaN- Not a Number

The prediction performance of 160 mg P kg^{-1} clearly was lower. For that reason it would make more sense to calculate the RMSEP based on the other P levels (0; 20; 40 and 80 mg P kg^{-1}). Thus, the general RMSEP for these four P levels in this soil type was 32.65 mg P kg^{-1}.

Also, for the Andosol the RMSEP values decreased from the P level 0 mg P kg^{-1} to 20 mg P kg^{-1}. In both cases the P levels added were over predicted on the regression model (Figure 5-10).

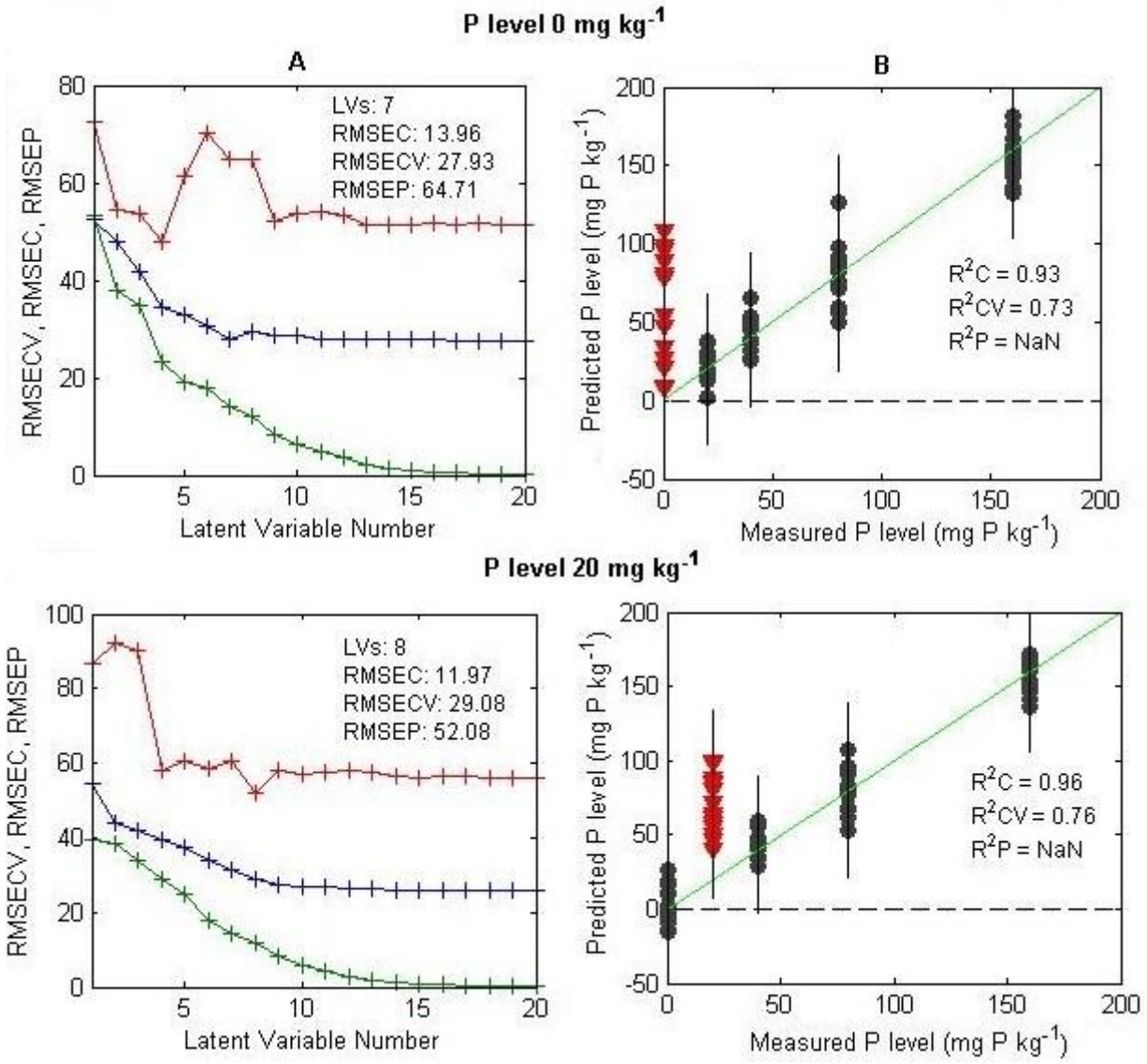


Figure 5-10. Vis/NIR prediction of different P levels (0 and 20 mg P kg^{-1}) added to the Andosol samples A- SIMPLS Variance captured and Statistics B- PLS regression model.

LVs- latent variables; RMSE- root mean square error; R^2-Coefficient of determination
C- Calibration; CV- Cross Validation; P- Prediction
NaN- Not a Number

The RMSEP values at 40 mg P kg^{-1} increased significantly with respect to the previous levels. Therefore, this value is over predicted in the regression model. The lowest value of RMSEP and the better prediction was obtained at the level of 80 mg P kg^{-1} (Figure 5-11).

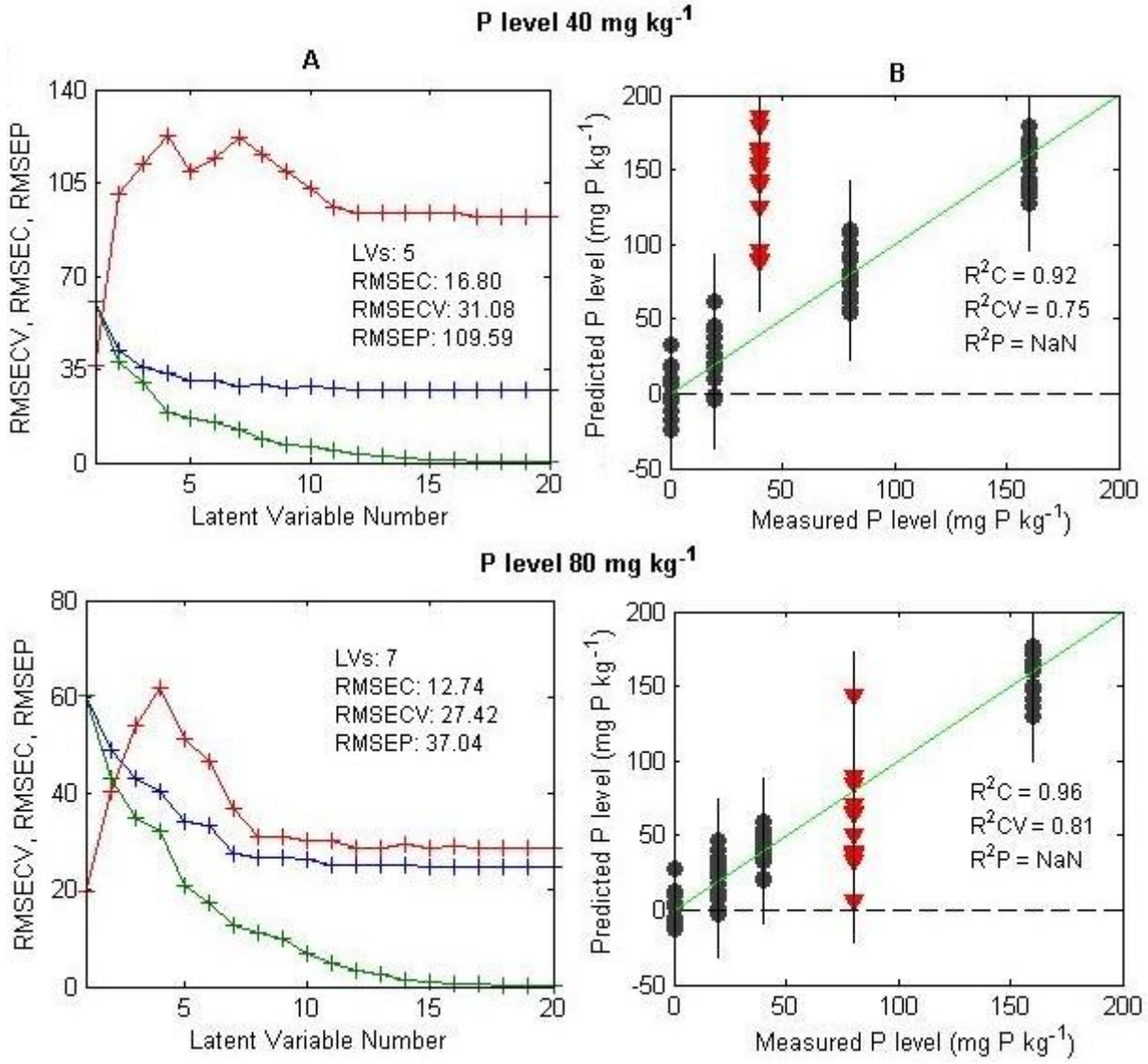


Figure 5-11. Vis/NIR prediction of different P levels (40 and 80 mg P kg^{-1}) added to the Andosol samples A- SIMPLS Variance captured and Statistics B- PLS regression model.

LVs- latent variables; RMSE- root mean square error; R^2-Coefficient of determination
C- Calibration; CV- Cross Validation; P- Prediction
NaN- Not a Number

Similar to the result observed in Cambisol, the RMSEP at P level of 160 mg P kg^{-1} showed a high difference with respect to the RMSECV. Also, this P level was under predicted by the regression model (Figure 5-12).

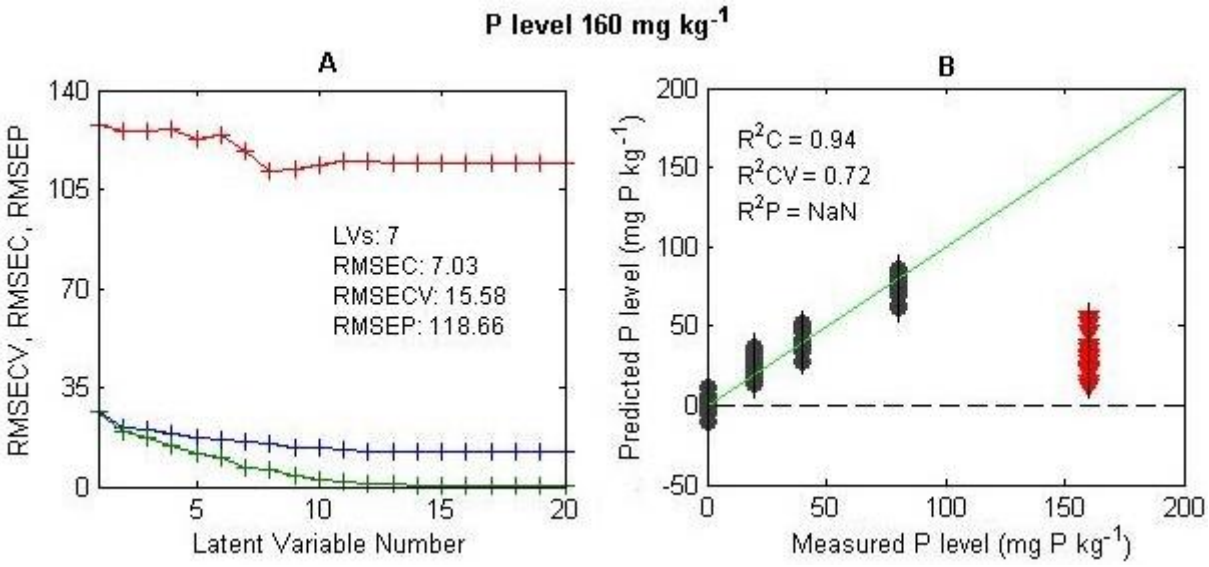


Figure 5-12. Vis/NIR prediction of a P level (160 mg P kg^{-1}) added to the Andosol samples A- SIMPLS Variance captured and Statistics B- PLS regression model.

LVs- latent variables; RMSE- root mean square error; R^2-Coefficient of determination
C- Calibration; CV- Cross Validation; P- Prediction
NaN- Not a Number

Based on the lower prediction observed at this level the general RMSEP was calculated with the other four P levels. The value of RMSEP obtained was 69.10 mg P kg^{-1}.

In coincidence with the observed in Cambisol and Andosol, the RMSEP value for the Ferralsol (Sega) decreased from the P level 0 mg P kg^{-1} to 20 mg P kg^{-1}. However, differ with the previous soil types regarded with the fact that at 20 mg P kg^{-1} the RMSEP was lower than the RMSECV (Figure 5-13).

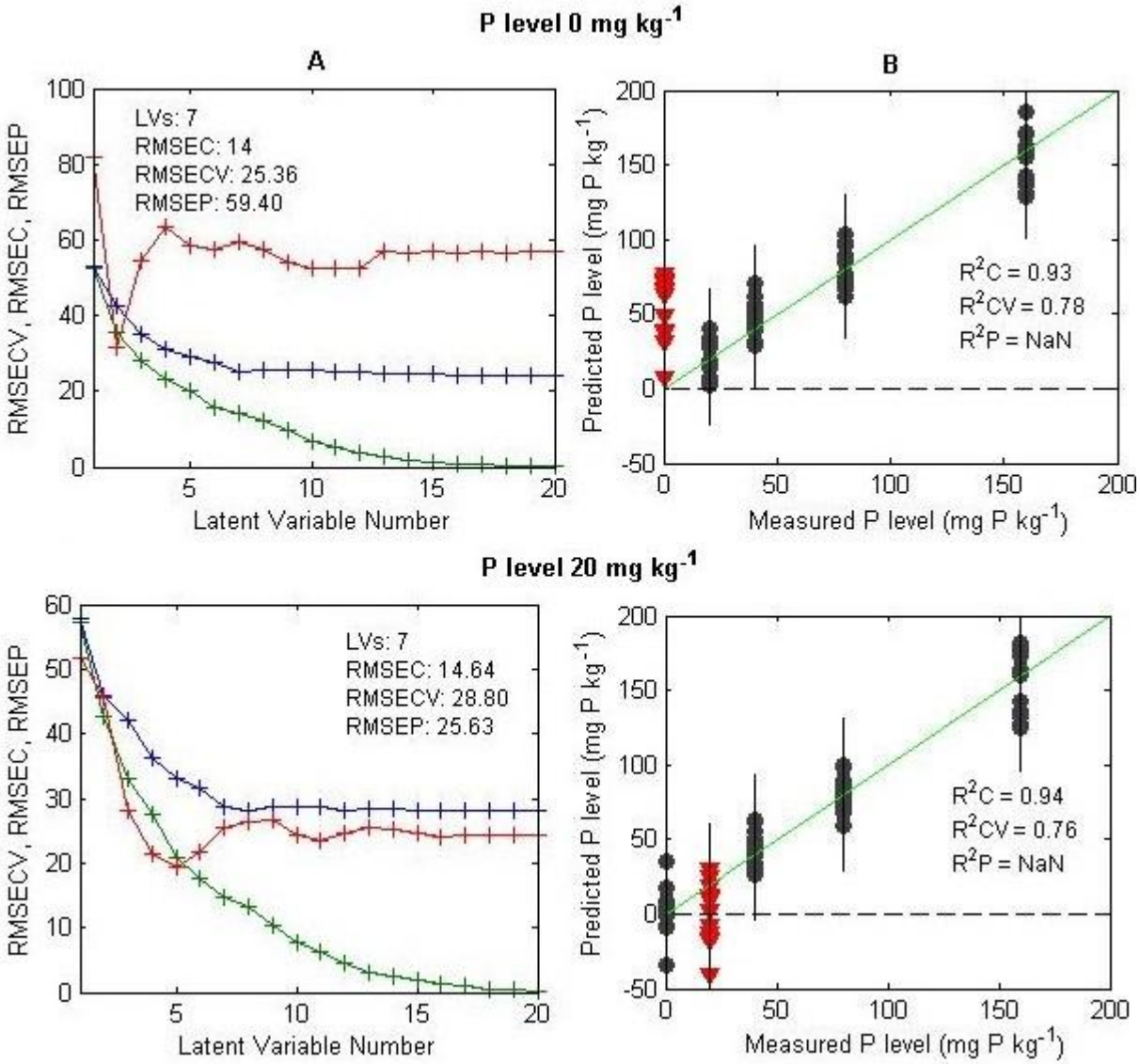


Figure 5-13. Vis/NIR prediction of different P levels (0 and 20 mg P kg^{-1}) added to the Ferralsol (Sega) samples. A- SIMPLS Variance captured and Statistics B- PLS regression model.

LVs- latent variables; RMSE- root mean square error; R^2-Coefficient of determination
C- Calibration; CV- Cross Validation; P- Prediction
NaN- Not a Number

In this soil type the RMSEP value at 40 mg P kg^{-1} increased with respect to the previous level, but also was higher than the RMSEP observed at the P level of 80 mg P kg^{-1}. The regression model trends to over predict the P level of 40 mg P kg^{-1} (Figure 5-14).

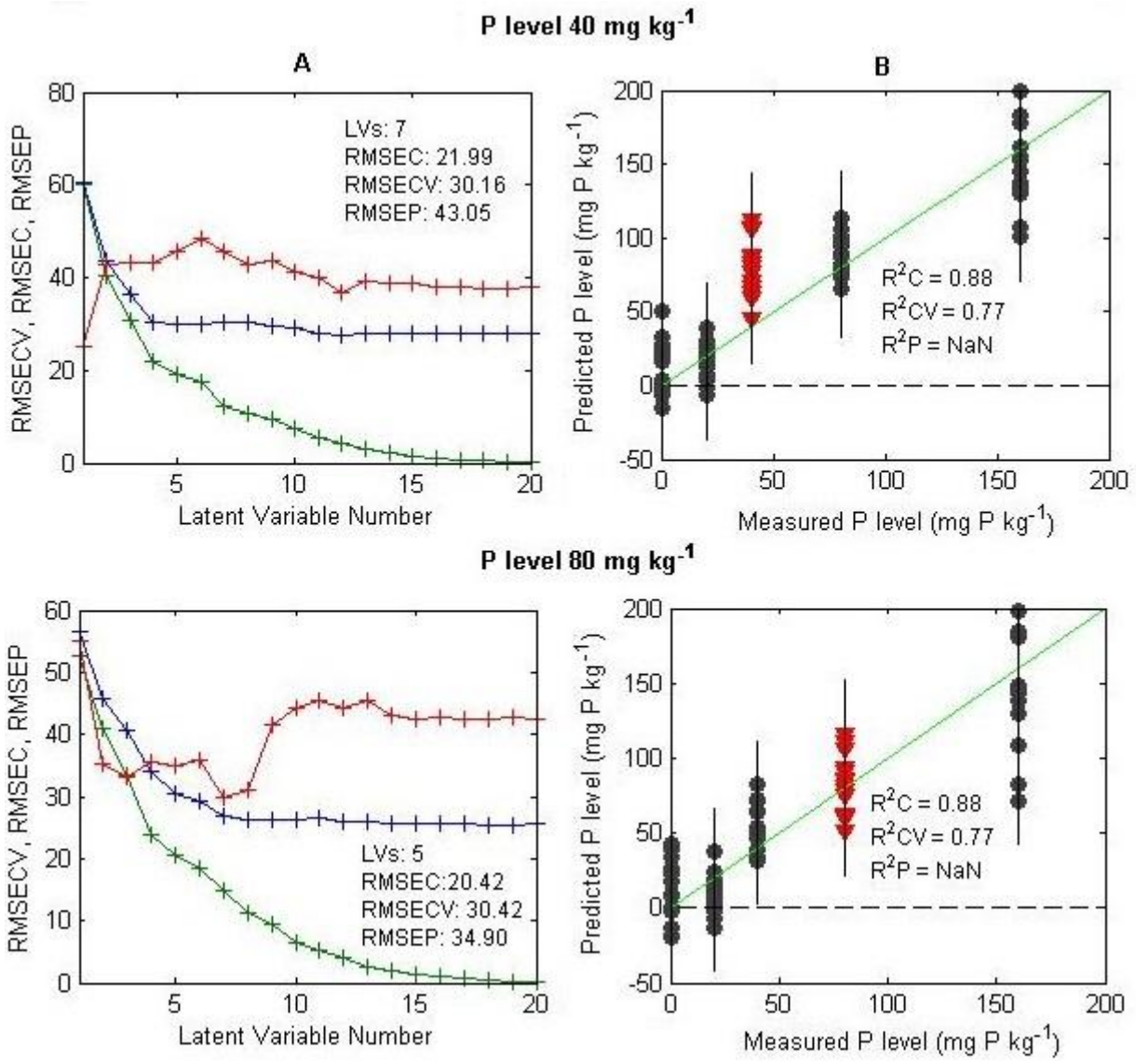


Figure 5-14. Vis/NIR prediction of different P levels (40 and 80 mg P kg^{-1}) added to the Ferralsol (Sega) samples. A- SIMPLS Variance captured and Statistics B- PLS regression model.

LVs- latent variables; RMSE- root mean square error; R^2-Coefficient of determination
C- Calibration; CV- Cross Validation; P- Prediction
NaN- Not a Number

The RMSEP at P level of 160 mg P kg^{-1} is keeping completely extrapolated like the others soil types previously analysed. Also this P level was totally under predicted on the regression model (Figure 5-15).

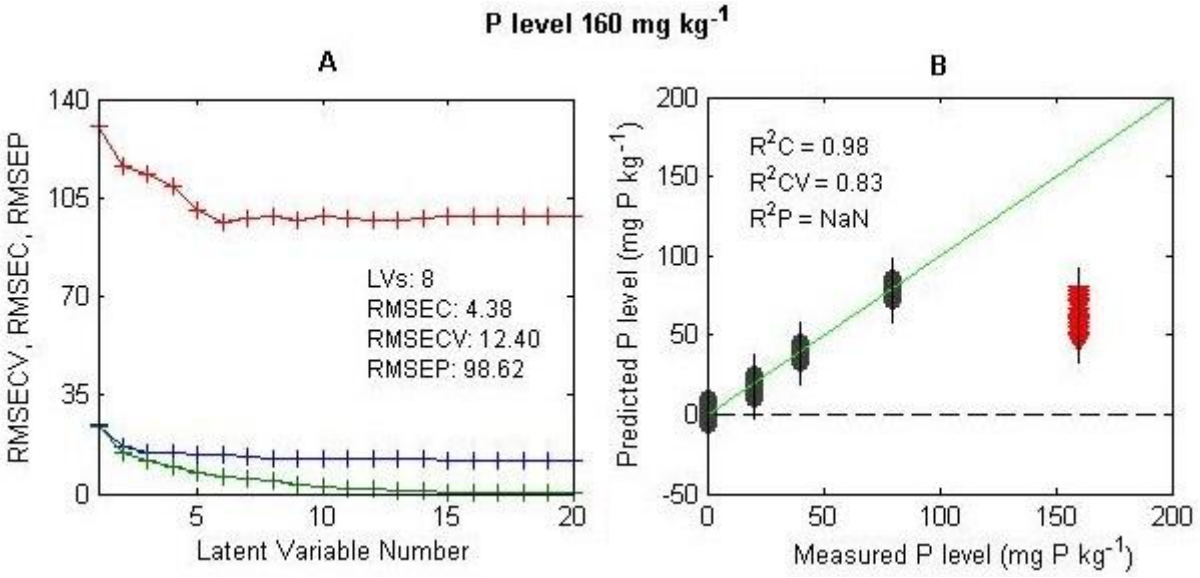


Figure 5-15. Vis/NIR prediction of different P levels (160 mg P kg^{-1}) added to the Ferralsol (Sega) samples A- SIMPLS Variance captured and Statistics B- PLS regression model.

LVs- latent variables; RMSE- root mean square error; R^2-Coefficient of determination
C- Calibration; CV- Cross Validation; P- Prediction
NaN- Not a Number

In line with the results obtained for Cambisol and Andosol the general RMSEP was calculated with four P levels (0; 20; 40; 80 mg P kg^{-1}). In this respect, a general RMSEP of 43.73 mg P kg^{-1} was obtained for the P levels analysed.

In comparison with the other soils previously analysed, the Ferralsol (Ivory) was the only soil type which the RMSEP value at 0 mg P kg^{-1} is over 100. In this case the regression model showed an over prediction. For 20 mg P kg^{-1} the RMSEP decreased and actually it was the level with the lowest value of this statistic in this type of soil (Figure 5-16).

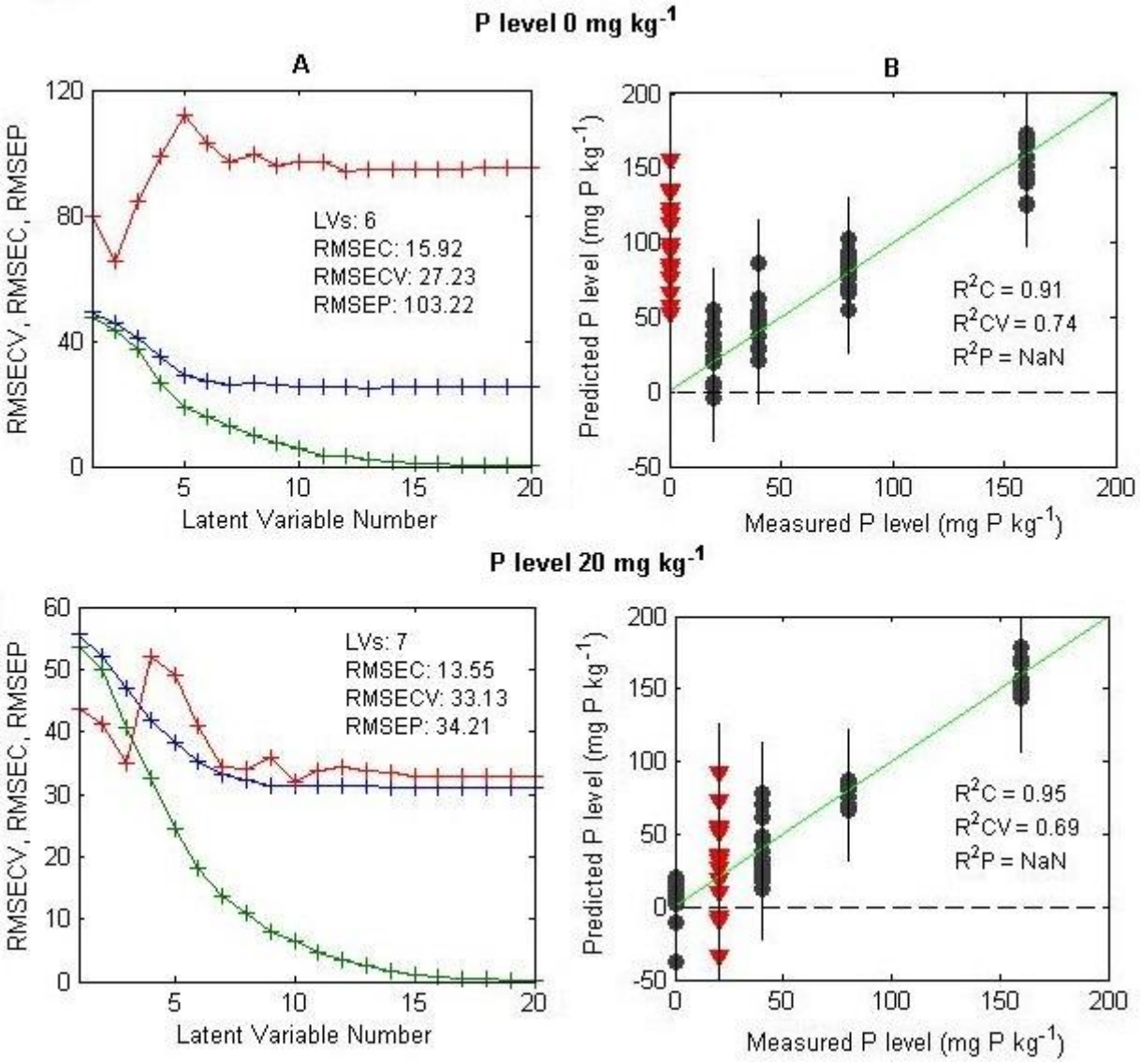


Figure 5-16. Vis/NIR prediction of different P levels (0 and 20 mg P kg^{-1}) added to the Ferralsol (Ivory) samples. A- SIMPLS Variance captured and Statistics B- PLS regression model.

LVs- latent variables; RMSE- root mean square error; R^2-Coefficient of determination
C- Calibration; CV- Cross Validation; P- Prediction
NaN- Not a Number

In comparison with the P level of 20 mg P kg^{-1}, the RMSEP values increased at 40 mg P kg^{-1} and at 80 mg P kg^{-1}. Also, both values from each level showed some similarities between them. However, the level of 80 mg P kg^{-1} is over predicted on the regression model (Figure 5-17).

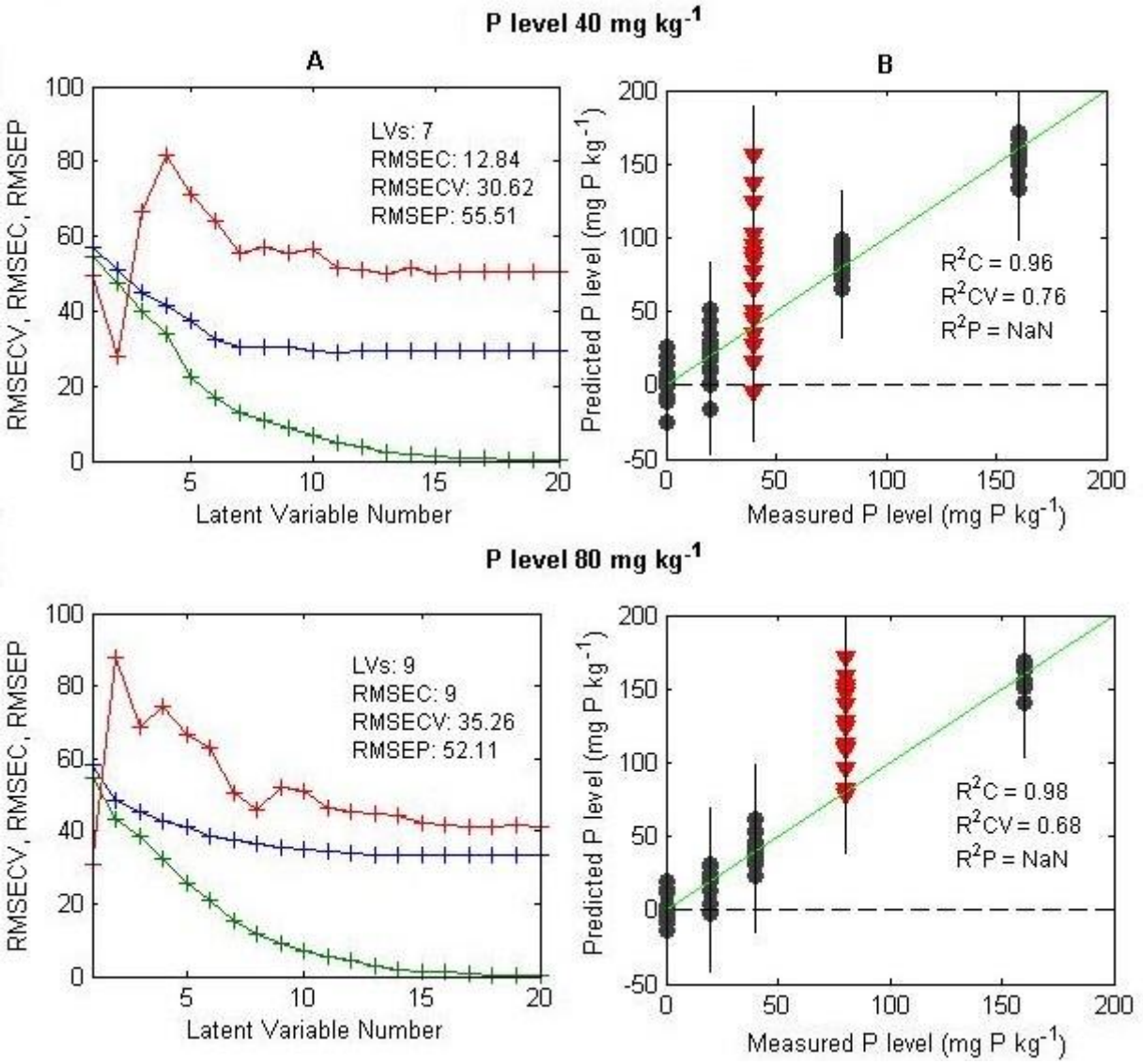


Figure 5-17. Vis/NIR prediction of different P levels (40 and 80 mg P kg^{-1}) added to the Ferralsol (Ivory) samples. A- SIMPLS Variance captured and Statistics B- PLS regression model.

LVs- latent variables; RMSE- root mean square error; R^2-Coefficient of determination
C- Calibration; CV- Cross Validation; P- Prediction
NaN- Not a Number

The RMSEP and the regression model at P level of 160 mg P kg^{-1} showed similar results previously observed on the others soil types analysed (Figure 5-18).

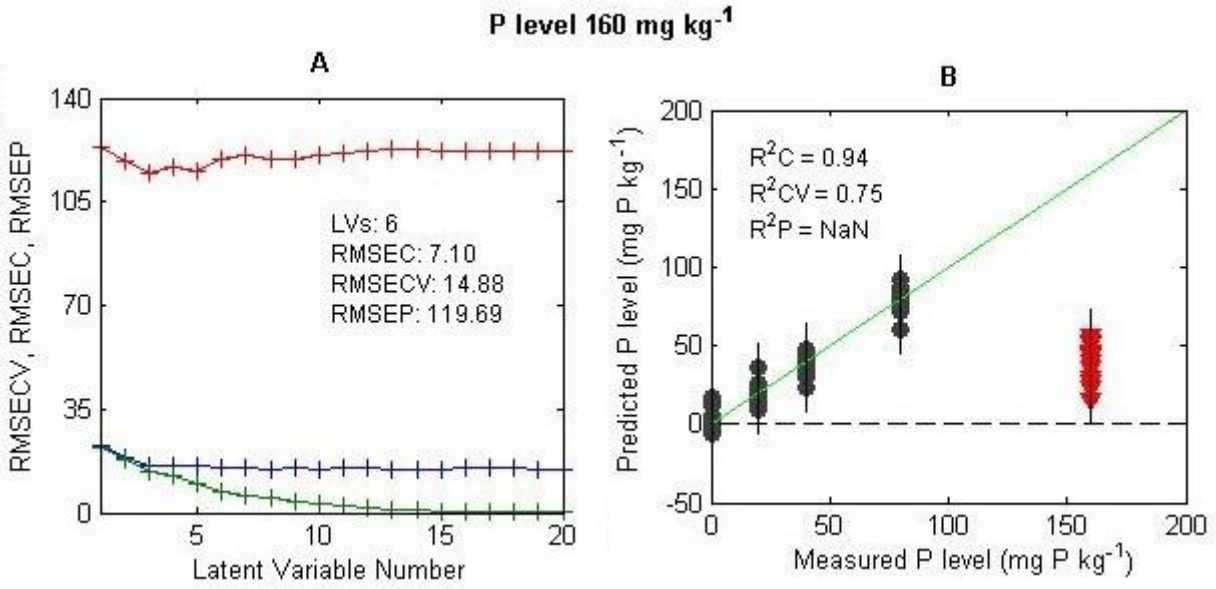


Figure 5-18. Vis/NIR prediction of a P level (160 mg P kg^{-1}) added to the Ferralsol (Ivory) samples A- SIMPLS Variance captured and Statistics B- PLS regression model.

LVs- latent variables; RMSE- root mean square error; R^2-Coefficient of determination
C- Calibration; CV- Cross Validation; P- Prediction
NaN- Not a Number

In this type of soil the prediction performance was lower at this P level of 160 mg P kg^{-1}. Therefore, only the other four P levels were used for calculating the general RMSEP. The RMSEP value was 64.75 mg P kg^{-1}.

Figure 5-19 illustrates that Ferralsol (Kuinet) is the only soil type which showed a lower RMSEP at 0 mg P kg^{-1} than the RMSEP of 20 mg P kg^{-1}. The regression model clearly over predicted the P level of 20 mg P kg^{-1}.

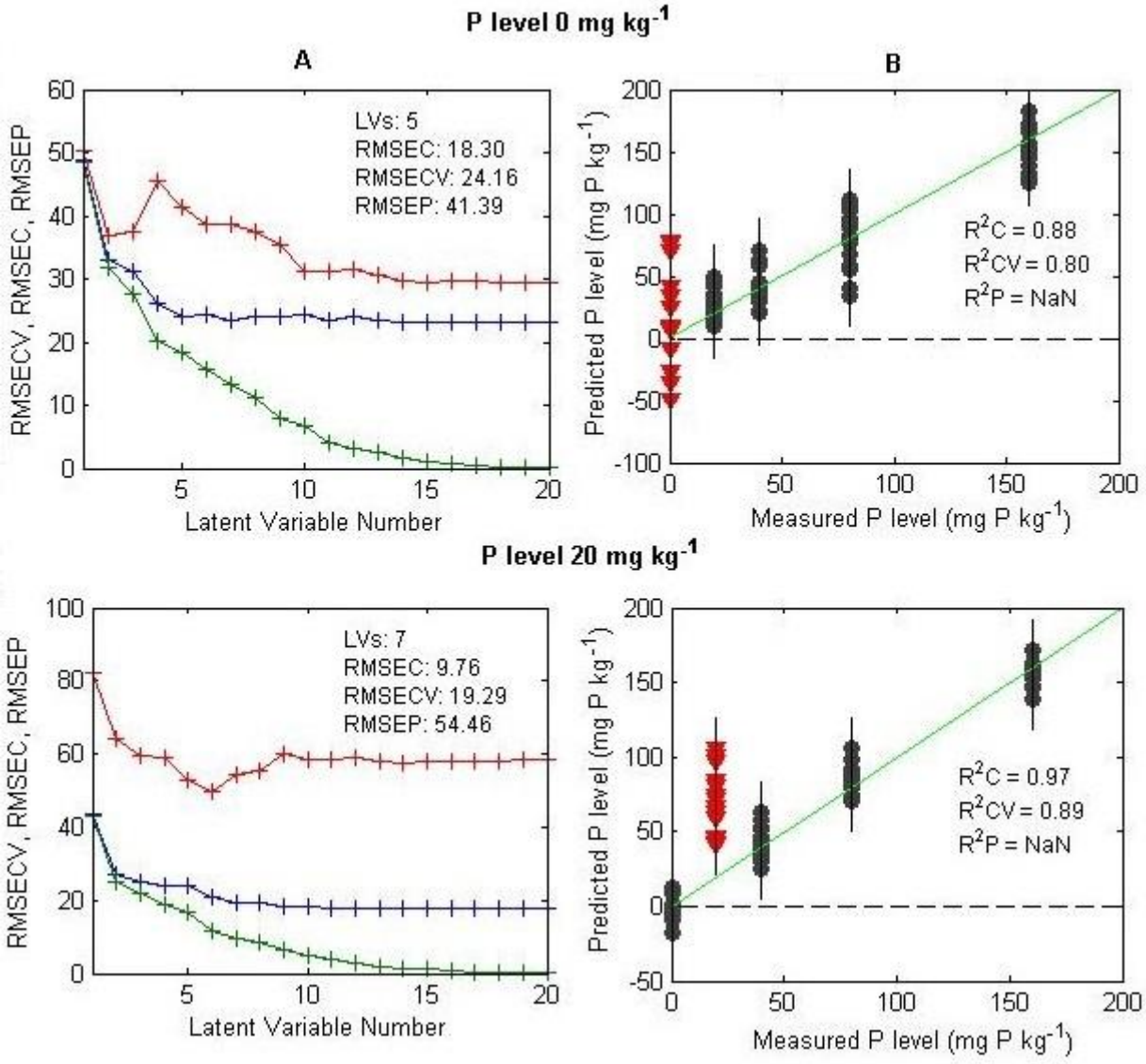


Figure 5-19. Vis/NIR prediction of different P levels (0 and 40 mg P kg^{-1}) added to the Ferralsol (Kuinet) samples. A- SIMPLS Variance captured and Statistics B- PLS regression model.

LVs- latent variables; RMSE- root mean square error; R^2-Coefficient of determination
C- Calibration; CV- Cross Validation; P- Prediction
NaN- Not a Number

The P levels of 40 mg P kg^{-1} to 80 mg P kg^{-1} predicted similarly with the lowest values of RMSEP observed for this soil type. However, both levels differed a little regarding accuracy. The performance of the regression model trends to be better at 80 mg P kg^{-1}(Figure 5-20).

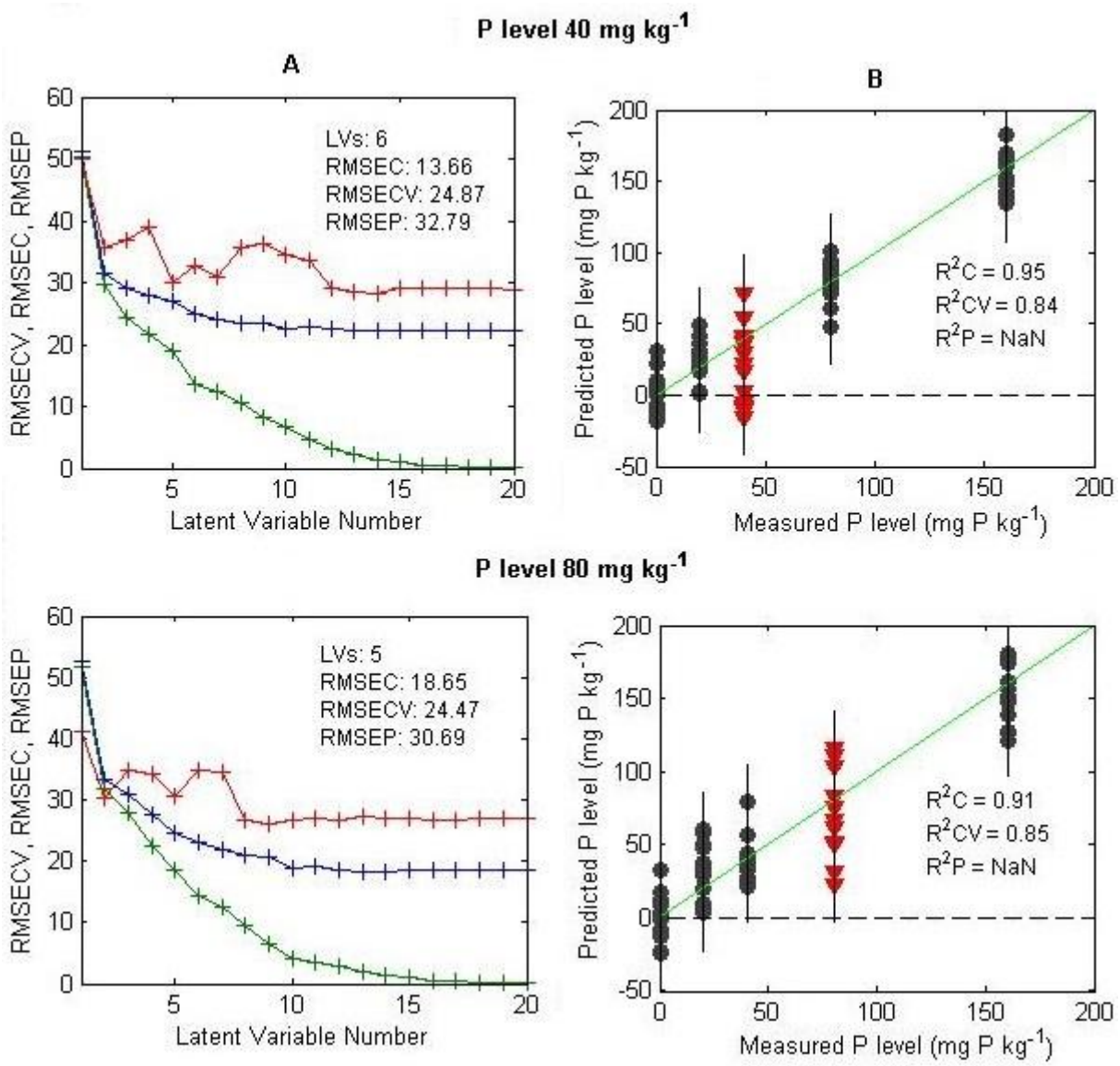


Figure 5-20. Vis/NIR prediction of different P levels (40 and 80 mg P kg^{-1}) added to the Ferralsol (Kuinet) samples. A- SIMPLS Variance captured and Statistics B- PLS regression model.

LVs- latent variables; RMSE- root mean square error; R^2-Coefficient of determination
C- Calibration; CV- Cross Validation; P- Prediction
NaN- Not a Number

Figure 5-21 shows a coincidence with the rest of the soil types analysed at the P level of 160 mg P kg^{-1}. This P level is under predicted too on the regression model. However, the RMSEP was lower than those observed in other soil types.

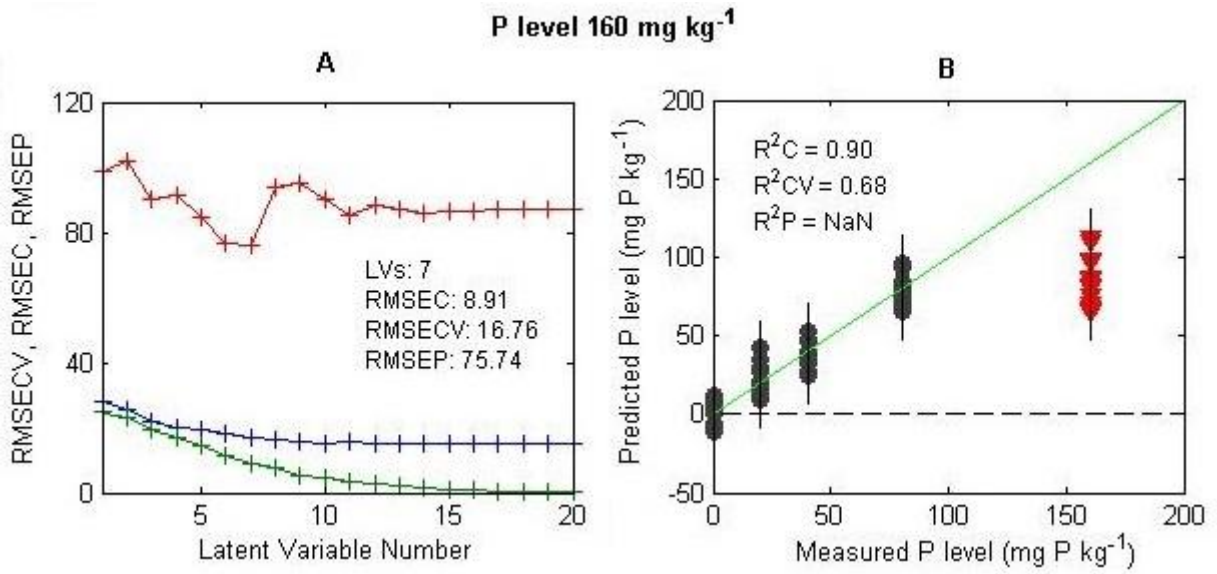


Figure 5-21. Vis/NIR prediction of a P level (160 mg P kg^{-1}) added to the Ferralsol (Kuinet) samples A- SIMPLS Variance captured and Statistics B- PLS regression model.

LVs- latent variables; RMSE- root mean square error; R^2-Coefficient of determination
C- Calibration; CV- Cross Validation; P- Prediction
NaN- Not a Number

The prediction performance for the Ferralsol (Kuinet) at this P level was lower than the obtained in the other four P levels. Therefore, the general RMSEP of 41.26 mg P kg^{-1} was calculated over the 0; 20; 40 and 80 mg P kg^{-1} analysed.

5.4. Conclusions

This is one of the first studies which used Vis/NIR spectroscopy to predict different P levels added to several soil samples from three types of soil located in five regions of the world. It was demonstrated that the pattern of soil reflectance signatures for each soil type was similar with increasing values from Vis to NIR region. The determination of soil spectral signatures after removing other soil constituents by acid leaching, might improve the prediction accuracy of soil P concentration. The samples from the same soil type but with different levels of P added, showed different magnitudes of reflectance spectra. Also, all types of soils showed water absorption bands at the same wavelengths. The RMSEP and the regression models indicated that the prediction of the added P levels was very inaccurate in the different soil types. In this sense, the effect of different P levels on the spectra showed that the RMSEP values decreased from 0 mg P kg^{-1} to 20 mg P kg^{-1} for all the soil types except for Ferralsol (Ivory). The P level 160 mg P kg^{-1} was completely extrapolated in all the soil types. The prediction of the added P levels in the Cambisol samples was better than in the other soil types. For 20, 40 and 80 mg P kg^{-1} added P the RMSEP values were comparable with their respective RMSECV. This suggests that the effect of the P level in Cambisols could be extracted from the Vis/NIR spectra independent of the OM content. However, it should be noted that the variation in the predictions for different spectra acquired for the same P level was large. This indicates that there is a lot of spectral variability which is not caused by the P level. This effect might be reduced by averaging more scans and/or preprocessing of the acquired spectra.

Chapter 6

Vis/NIR measurement of the average fertility values for soil samples at landscape level

6.1. Introduction

A soil landscape is an area of land with a common suite of soil types and landscape attributes (Murphy et al., 2001). According to Valentine et al. (1978) a *soil landscape* is thought of as the total ecosystem with which a particular soil is associated, with emphasis placed on the soil itself. It could be used at any level of the soil classification. Many soil landscapes occur in more than one physiographic region. For instance Murphy et al. (2001) considered that the mapping of soil landscapes through the investigation and collection of detailed soil and landscape profiles and the analyses of fertility parameters provides a powerful natural resource database on which to make informed environmental decisions.

The productivity of a soil is determined by its properties, which are the result of its development. The development of a soil is largely determined by its position in the landscape. The variability of soil fertility parameters across farm landscapes impacts the land management, the crop production potential as well as the soil and water conservation (Veseth, 1986). Therefore, characterizing the spatial heterogeneity in the physical (texture) and chemical (nutrient content) properties across the agricultural fields is important to identify the scale of spatial variation of these soil fertility parameters. This characterization contributes to a more efficient and effective fertiliser management and has been performed all over the world. However, according to Moore et al. (1993) and Florinsky et al. (2002) conventional soil surveys (sampling and soil fertility analyses) are not sufficient to obtain all the quantitative information about the spatial distribution of soil fertility parameters that would be needed for optimal fertiliser use. Thus, in most developed countries soil sampling and wet chemical analyses are regularly performed to determine the average soil fertility parameters for an agricultural field. These analysis methods require that moist soil samples collected from the field must be air dried as soon as possible before sending to a soil testing

laboratory. After the whole process of sieving and grinding the soil samples, wet chemical analyses are performed on them. Then, the final results are communicated to the farmers for fertiliser management decision-making. Recently, several researchers such as Vasques et al. (2009), Stevens et al.(2008), Bogrekci and Lee (2005a) have reported on the potential of Vis/NIR spectroscopy to quantify different soil fertility parameters and obtained the best results on air-dried soil samples. The potential of Vis/NIR spectroscopy for measuring the most important soil fertility parameters in Cuban agricultural soils has been evaluated in this study. The materials and methods are presented in section 6.2 and the results and discussion are presented in section 6.3. Finally, conclusions are drawn from this study in section 6.4.

6.2. Materials and methods

The setup for acquiring Vis/NIR diffuse reflectance spectra of the air-dried soil samples in the laboratory consisted of a diode array spectrophotometer (CORONA PLUS REMOTE Vis/NIR SB, Zeiss, Jena, Germany) and an OMK 500-H measuring head connected to it with an optical fibre-bundle. The Aspect –*plus* software supplied by Carl Zeiss Jena GmbH was used to control the spectrophotometer and collect soil spectra. The spectra were collected from the air-dried soil samples in the Vis/NIR region (399 – 1697 nm) as following the procedure described in Chapter 4.

Prediction models linking the acquired Vis/NIR reflectance spectra to the selected soil fertility parameters were built using the linear multivariate regression method Partial Least Squares (PLS), and the nonlinear multivariate regression methods Support Vector Machines (SVM) and Locally Weighted Regression (LWR). The PLS regression relates the variations in one response variable to the variations of several predictors (wavelengths) as stated by Yang et al. (2011). This method is based on orthogonal transformation technique. It reduces the complexity of modelling and eliminates the adverse effects of multicollinearity among spectral variables. SVM focuses on minimising a bound on the risk function, rather than minimising the error in training data. In this way, the over-fitting problem is prevented (Karimi et al., 2008). For nonlinear cases, SVM uses a so-called kernel technique to plot the data into a higher dimensional feature space, where linear functions can be applied. More detailed information on SVM may be found in Vapnik (1995), Vapnik et al. (1997), Smola and Scholkopf (1998). In LWR, the spectra is first compressed by using principal component analysis, and then the Mahalanobis

distance is computed on the first principal components which accounts for a given percentage of cumulative explained variance. After this procedure local PLS calibrations for each unknown sample are carried out in the spectral space using its k-nearest neighbours which are weighted according to their distance from the unknown sample (Ramirez-Lopez et al., 2013).

For the Vis/NIR prediction at landscape level, the 189 samples taken from the different Cambisol fields in the Villa Clara province (Chapter 3) were randomly divided into a calibration set of 126 samples and a validation set of 63 samples. Similarly, the 144 Vertisol soil samples were randomly split into a calibration set of 108 samples and and a validation set of 36 samples.

During the analyses several samples were identified as outliers, or strange values whose presence could alter the results in a remarkable way. The criterion for identifying an outlier was based on the examination of the data for unusual observations which do not conform to the pattern established by other observations. Outlier analysis was also performed by screening from the soil data sets those data for which measured concentrations of soil fertility parameters were significantly higher or lower than the sample population and therefore could have been influenced by the sampling and chemical analysis procedures. After inspection of the spectra these outliers were removed from the dataset.

All calculations were performed in MATLAB 7.9 (R2009b, The Mathworks, Nattick, MT). The Venetian blinds cross-validation strategy (10 splits) was used in this research to evaluate the prediction error as a function of the model complexity (Leung, 2005).

In order to select the most suitable pre-processing methods for the Cambisol and Vertisol spectra, a trial and error process was followed. The choice of these methods potentially has a crucial impact on the interpretation of the results. For this reason, different data preprocessing methods commonly used for processing soil spectral data sets were test. Depending on the combination of preprocessing methods differences regarding to several statistics were observed. Therefore, the selection of the best pre-processing method was related to the prediction performance of the calibration models. This performance was evaluated based on the calibration statistics: coefficient of determination R^2 in cross-validation and test set prediction, root mean square error of cross validation RMSECV and test set prediction RMSEP. Also, the statistics R^2 in test set prediction, RPD (ratio of

performance deviation) and RER (ratio of error range) were used for comparing the calibration accuracies obtained for the different soil fertility parameters. The RPD is a statistic calculated by dividing the standard deviation (SD) of reference data used in the validation set by standard error of prediction (SEP):

$$RPD = \frac{SD}{SEP} \qquad (6.2.1)$$

The RER is computed by dividing the range in the reference data used in the validation set by the SEP:

$$RER = \frac{Range}{SEP} \qquad (6.2.2)$$

Calibration accuracy was assessed using the guidelines proposed by Malley et al. (2004) and Nduwamungu et al. (2009) for environmental samples like soil. These criteria are summarised in Table 6-1.

Table 6-1

Guidelines for assessing the calibration accuracy.

Calibration accuracy	R^2 [a]	RPD [b]	RER [c]	Source
Excellent	$R^2 > 0.95$	RPD >4	RER > 20	Malley et al. (2004)
Successful	$0.90 \leq R^2 \leq 0.95$	$3 \leq RPD \leq 4$	$15 \leq RER \leq 20$	Malley et al. (2004)
Moderately successful	$0.80 \leq R^2 < 0.90$	$2.25 \leq RPD < 3$	$10 \leq RER < 15$	Malley et al. (2004)
Moderately useful	$0.70 \leq R^2 < 0.80$	$1.75 \leq RPD < 2.25$	$8 \leq RER < 10$	Malley et al. (2004)
Less reliable	$R^2 < 0.70$	RPD < 1.75	RER < 8	Nduwamungu et al. (2009)

[a] R^2- coefficient of determination; [b] RPD- ratio of performance deviation; [c] RER- ratio of error range

Data were analysed with Forward Interval PLS model (iPLS) to determine those specific wavelength regions important for better estimating each soil fertility parameter. This method is based on the division of the full spectrum into smaller and equidistant intervals (15) of equal width (28 variables). The RMSECV is calculated for each interval and compared with the value obtained for the full spectrum model. Regions that present the smallest value of RMSECV are then chosen (Müller et al., 2011; Nørgaard et al., 2000). The calculations were performed in MATLAB 7.9.

6.3. Results and discussion

6.3.1. Prediction of soil fertility parameters in Cambisol

The final selected pre-processing method included logarithm transformation [Log (1/R)], smoothing and dataset centering (Mean Centre). Spectral pre-processing with mathematical functions is commonly used to correct for non linearities and electronic noise of the detector. Also, pre-processing methods aim to remove variation in the spectra which is not caused by the component of interest, but due to light scattering and chemical interference. In figure 6-1 the diffuse reflectance spectra of Cambisol at landscape scale (calibration and validation set) before and after pre-processing are shown.

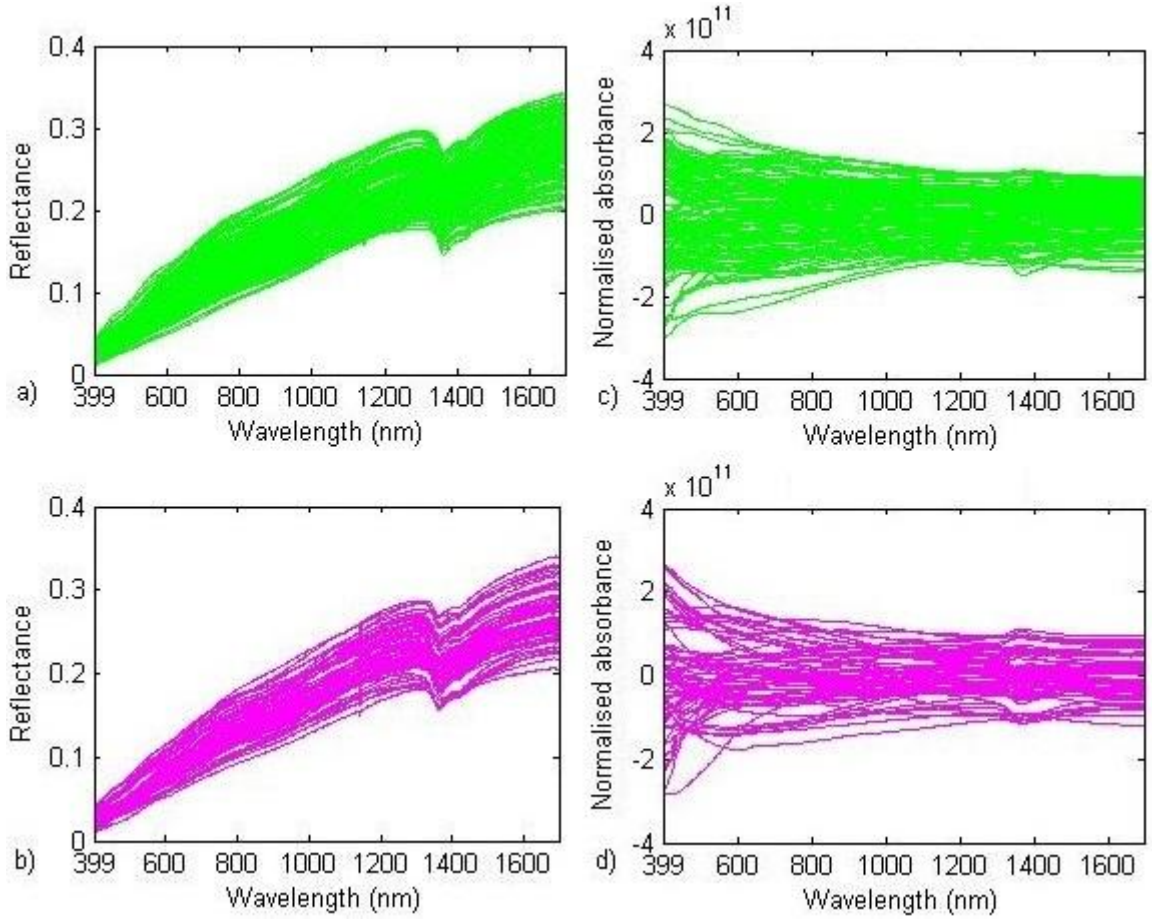


Figure 6-1. Soil spectral reflectance before and after preprocessing for measuring the average fertility at landscape level on Cambisol soil. (a) Raw data of calibration, (b) Preprocessed data of calibration, (c) Raw data of validation (d) Preprocessed data of validation.

These curves (left) are a graphical representation of the spectral reflectance of the air-dried soil samples used in this research, as a function of the studied wavelength range in the Vis/NIR region. In the visible region (380 – 780 nm) the reflectance is lower than in the NIR region (780 – 1697 nm). Also, around the wavelength of 1400 nm, the

reflectance values show a local minimum. The absorption of light observed in the 1350 to 1450 nm region is primarily by O-H, C-H and N-H bonds. Stenberg et al. (2010) explained that the absorptions in the NIR region result from the overtones of OH, SO_4, and CO_3 groups, as well as combinations of fundamental features of H_2O and CO_2.

The performance of the selected models is illustrated in Figure 6-2, which shows the correlations between the measured values and those predicted by the Vis/NIR model. In cross validation, the best prediction coefficient (R^2) was obtained for OM (0.93) with LWR regression analysis.

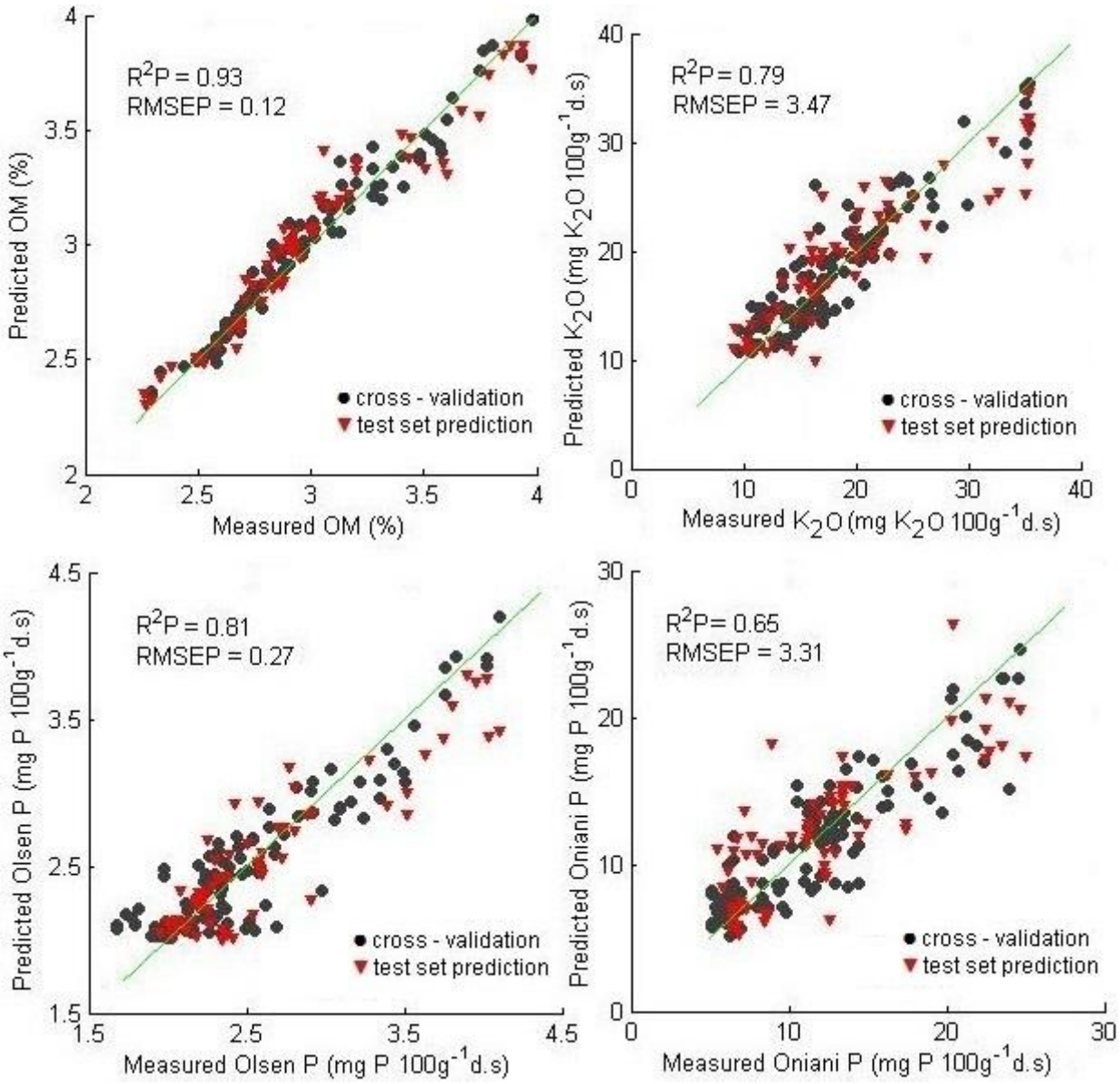


Figure 6-2. Vis/NIR predicted versus chemical conventional analyses of soil fertility parameters at landscape level on Cambisol soil; OM (LWR), K_2O (LWR), Olsen P (SVM), Oniani P (SVM).

Also, for K_2O this coefficient (0.79) was better with the same nonlinear regression model (LWR). For OM the data points were closely grouped around the target line both in the calibration set and in the prediction set, while they were slightly more scattered around this line in both sets for Olsen P and Oniani P; for K_2O a higher scattering was observed. By comparing the SVM prediction models for Olsen P and Oniani P, it was observed that the prediction performance for Oniani P was considerably worse (R^2 of 0.81 vs. 0.65). This suggests that the P content measured by Vis/NIR spectroscopy corresponds better to the P content assessed with the Olsen method than with the Oniani method. A possible explanation for this may be found in the fact that the Olsen method for determining the available P content shows a better performance in calcareous and neutral soils such as Cambisols. The Oniani method, when used in calcareous soils for analysing the available P content, tends to dissolve calcium carbonate ($CaCO_3$). This method is a good descriptor of available phosphorus in soils; but according to Herrera et al. (2002) this method only provides a static[7] estimation of plant-available P. From these results it could be derived that laboratory methods might influence the prediction ability of P in Cambisol by means of Vis/NIR spectroscopy technique.

Table 6-2 shows the results of the better regression models and others used for the prediction of the soil fertility parameters at landscape level. The three regression models (PLS, SVM, LWR) developed for OM achieve a successful prediction performance. However, LWR produces more accurate results with the highest R^2 of prediction (0.93) and the lowest RMSEP (0.12). Also, LWR model shows a minimum difference (0.01) between RMSEC and RMSECV; however there was not difference between these statistics for the other two regression models. The higher accuracy obtained for OM was based on the characteristic of LWR in selecting the most suitable soil samples for this soil fertility parameter. In line with Stenberg et al. (2010), OM is one of the fundamental constituents of the soil which has well-recognisable absorption features in the Vis/NIR region. This indicates that the prediction of this constituent in new Cambisol samples based on the Vis/NIR spectra will be reliable. Similar statistical results for K_2O were obtained from PLS, LWR and SVM. In all cases the prediction accuracy is classified as moderately useful ($0.70 \leq R^2 < 0.80$). This soil fertility parameter is not spectrally active in the Vis/NIR range.

[7] The Oniani method, as other chemical extraction methods to estimate plant-available soil P, extracts only some forms of P depending on the nature of the extraction solution; therefore the results obtained from this method not express the accurate amount of available P for plants in soil.

Table 6-2

Comparison between the best predictive models with others calibrated at landscape level on Cambisol soil.

Soil fertility parameter	Regression model	R^2 C [a]	R^2 CV [b]	R^2 P [c]	RMSEC [d]	RMSECV [e]	RMSEP [f]	RPD [g]	RER [h]	C Bias [i]	CV Bias [j]	P Bias [k]
	PLS	0.93	0.93	0.90	0.10	0.10	0.14	3.29	12.32	-0.00	-0.00	0.01
OM	**LWR**	**0.96**	**0.95**	**0.93**	**0.07**	**0.08**	**0.12**	**3.96**	**14.80**	**0.02**	**0.02**	**0.03**
	SVM	0.96	0.95	0.91	0.08	0.08	0.14	3.32	12.41	-0.00	-0.00	0.02
	PLS	0.77	0.75	0.77	3.00	3.12	3.62	2.08	7.26	-1.06	-0.01	0.03
K_2O	**LWR**	**0.83**	**0.79**	**0.79**	**2.72**	**3.07**	**3.47**	**2.79**	**9.74**	**0.48**	**0.51**	**2.18**
	SVM	0.83	0.78	0.78	2.66	3.02	3.49	2.17	7.57	-0.15	0.04	0.33
	PLS	0.71	0.69	0.69	0.29	0.31	0.35	1.82	6.21	-0.00	-0.00	-0.05
Olsen P	LWR	0.83	0.78	0.81	0.23	0.26	0.27	2.35	8.01	0.03	0.04	-0.03
	SVM	**0.83**	**0.81**	**0.81**	**0.22**	**0.24**	**0.27**	**2.37**	**8.10**	**0.00**	**-0.00**	**-0.05**
	PLS	0.75	0.73	0.64	2.57	2.65	3.54	1.60	5.56	-1.06	-0.01	0.39
Oniani P	LWR	0.80	0.74	0.65	2.33	2.62	3.61	1.61	5.61	0.37	0.40	0.93
	SVM	**0.79**	**0.73**	**0.65**	**2.37**	**2.65**	**3.31**	**1.70**	**5.91**	**-0.21**	**-0.12**	**0.03**

[a] R^2 C- coefficient of determination of calibration; [b] R^2 CV- coefficient of determination of cross validation; [c] R^2 P- coefficient of determination of prediction
[d] RMSEC- root mean square error of calibration; [e] RMSECV- root mean square error of cross validation; [f] RMSEP- root mean square error of prediction;
[g] RPD- ratio of performance deviation; [h] RER- ratio of error range;
[i] C Bias- systematic deviation of calibration; [j] CV Bias- systematic deviation of cross validation; [k] P Bias- systematic deviation of prediction
Data in bold- Best regression model based on the R^2, RMSECV, RMSEP, RPD, RER

Therefore K_2O was estimated indirectly due to its high correlation (r=0.78) to a more spectrally active parameter such as OM, which influences most soil chemical properties. The soil OM can be directly related to the absorption in Vis/NIR spectra through a number of functional groups such as NH, CH, and CO groups. In this sense Irons et al. (1989) considered that an increase in the OM content of a soil generally causes a decrease of reflectance over the entire spectrum. A high OM content and hence, a strong decrease of overall reflectance, might even mask other absorption features in the soil spectra.

Olsen P was better predicted with SVM (R^2= 0.81). As the R^2 value in prediction was equal to that obtained in cross-validation, this model is expected to be quite robust. For this fertility parameter the prediction performance of the PLS regression model (R^2 of 0.69) was considerably worse than that for the other two regression models (LWR and SVM, both with R^2 of 0.81). This suggests that the relation between the preprocessed absorbance spectra and the Olsen P is nonlinear. Thus, SVM and LWR models are more reliable for the prediction of nonlinear data with a moderately successful accuracy for this fertility parameter.

Also, according to Borggaard (2001), a potential drawback with using PLS regression and all linear regression techniques is the possible nonlinearity of correlations between NIR spectra and the property of interest. This fertility parameter, just like K_2O, is not spectrally active in the Vis/NIR range. Therefore, the prediction accuracy for Olsen P could be substantially improved because this parameter is also highly correlated (r=0.89) with OM. In the case of Oniani P, all the models showed similar R^2 to each other ($0.64 \leq R^2 \leq 0.65$). According to the criteria suggested by Malley et al. (2004), this prediction accuracy is classified as less reliable. The prediction accuracy for P is considerably lower for Oniani than for Olsen method. The differences between both analytical methods resulted in different proportions of this soil fertility parameter.

In general for all soil fertility parameters, equally good or better prediction performance was obtained with the nonlinear regression models (LWR and SVM) than with the linear regression model (PLS). However, as already mentioned for OM also successful prediction results were obtained with PLS. This parameter showed a nonlinear relation with the others. Janik and Skjemstad (1995) suggested that nonlinear relationships between spectra and soil variables often occur with data covering wide ranges and can be the result of distortions of strong signals, and different mineralogical values with high and low

soil variable values in the calibration data sets. These results were in line with those obtained by Shao and He (2011). These authors reported R^2 equal to 0.82 and 0.80 for phosphorus and potassium respectively. In that case they used the Bray and Kurtz method for analysing phosphorus and flame atomic emission spectrometry for potassium.

As the K and P content were found to be highly correlated to the OM content, the prediction of the P and K content in soil from Vis/NIR spectra may be based on their correlation with the OM content. This was also concluded from the designed experiment described in chapter 5 where different levels of P added to the same soils could not be predicted from the corresponding Vis/NIR spectra.

6.3.2. Identification of effective wavelength intervals in Cambisol

The results of the above regression models were based on the full spectrum, which included a large number of wavelengths. Therefore, the most practical approach to simplify this methodology is to select the wavelength intervals with the lowest RMSECV that contribute most to the development of the Vis/NIR regression models.

The number of wavelengths required was reduced considerably to a maximum of 84 (each interval includes 28 wavelengths), which represent 20% of the full spectrum. Thus, different wavelength intervals were selected as important for each soil fertility parameter at landscape level, but some intervals were common in some of these parameters (Figure 6-3).

The interval number 1 (399 – 489 nm) was common to all the soil fertility parameters at landscape level. In addition, the interval number 2 (489 – 580 nm) and the interval number 3 (580 – 674 nm) were selected for K_2O and Olsen P. All these intervals fit in the Vis region (380 – 780 nm). In general, the absorptions in this region are mainly related to those minerals which contain iron.

Also, in other cases soil OM has a tendency to show broad absorption peaks in the Vis region, which are dominated by chromophores (spectrally active groups e.g. Fe, OH- in water and minerals, CO_3^{2-}, Al^{2+}, SO_4^{2-} in minerals). This statement is in agreement with Clark et al. (1990) and Clark (1999). The other two intervals selected for OM (1320 – 1402 nm; 1480 – 1555 nm) are included in the NIR region (780 – 2500 nm).

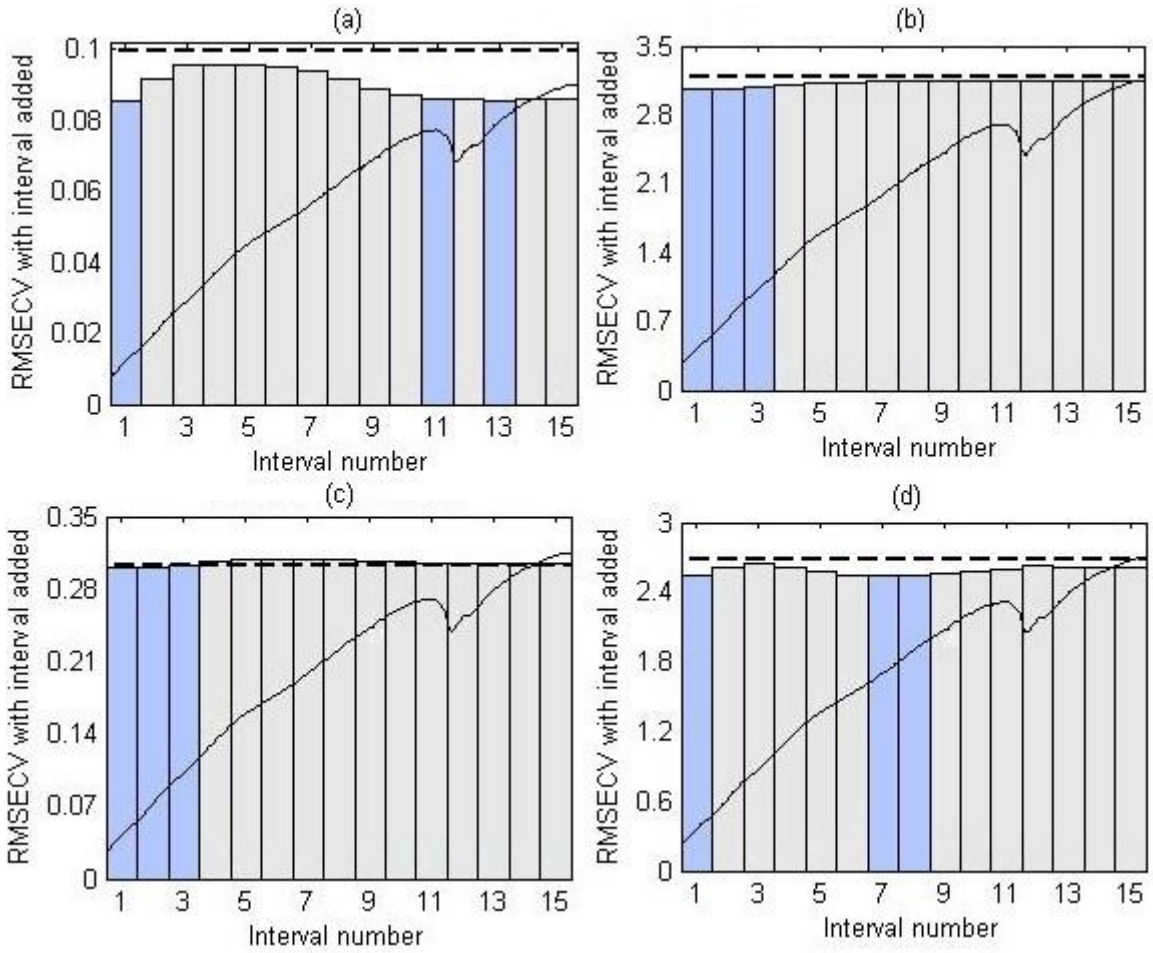


Figure 6-3. Spectral region selected (dark colour) by Forward Interval PLS model (iPLS) at landscape level on Cambisol soil.

Intervals selected:

(a) OM- [1 (399 – 489 nm), 11 (1320 – 1402 nm), 13 (1480 – 1555 nm)]
(b) K2O- [1 (399 – 489 nm), 2 (489 – 580 nm), 3 (580 – 674 nm)]
(c) Olsen P- [1 (399 – 489 nm), 2 (489 – 580 nm), 3 (580 – 674 nm)]
(d) Oniani P- [1 (399 – 489 nm), 7 (952 – 1050 nm), 8 (1050 – 1145 nm)]

Wavelengths included in the intervals number 1 and 11 selected for OM are in line with the results reported by several authors. For instance, Lee et al. (2003) estimated chemical properties in Florida and found that the correlation coefficients between the OM content and spectral reflectance in 428 nm and 1376 nm were higher. Mapping soil OM in the north-west part of Semnan province (Iran), Nowkandeh et al. (2013) selected 477 nm, as one of the best wavelengths for regression modelling. In Shan Dong province (China), Wang et al. (2013) reported that two of the optimal wavelengths which have the best fitness for predicting OM were 399 nm and 449 nm.

Also, Yu et al. (2013) reported for OM content that the wavelength of 492 nm and 1317 nm possessed the best prediction accuracy. Both wavelengths are nearest of 489 nm (interval number 1) and 1320 nm (interval number 2) respectively.

On the other hand, wavelengths included in the intervals selected in this study for K_2O, Olsen P and Oniani P have been reported by other researchers. In research carried out in two agricultural fields in north-eastern Mississippi, Thomasson et al. (2001) selected the wavelengths 425 nm, 525 nm, 575 nm, 625 nm and 675 nm for potassium (K). The last wavelength is close to 674 nm, which is the boundary of the interval number 3 of this fertility parameter. Also, for phosphorus (P) these authors selected the wavelengths 475 nm, 525 nm, 575 nm, 625 nm, 1075 nm, 1125 nm, 1225 nm, 1425 nm, 1475 nm and 1525 nm. For K, Lee et al. (2003) reported the wavelengths 428 nm, 444 nm and 522 nm. While for P the same authors reported the wavelengths 428 nm, 430 nm, 522 nm, 602 nm, 612 nm and 1100 nm.

Table 6-3 shows the results of the predictive regression models at landscape level based on the selected wavelength intervals. The three regression models (PLS, SVM, LWR) developed for OM achieve an excellent prediction performance. It was a better result than the successful accuracy obtained with the full spectrum. Also, LWR model gave the more accurate results with the highest R^2 of prediction (0.97) and the lowest RMSEP (0.08). The other soil fertility parameters also increased the R^2 of prediction and kept the model selected as the best with the full spectrum (LWR – K_2O; SVM – Olsen P and Oniani P). However, only K_2O and Oniani P differ as to the category related to the accuracy of performance observed with the full spectrum. The new accuracy of performance for K_2O achieves R^2 of 0.82 (moderately successful) and for Oniani P R^2 of 0.70 (moderately useful). The results demonstrate that all models developed for the soil fertility parameters using just 20% of the wavelengths performed better than with the full spectrum. Therefore, the highest accuracy levels and the lowest RMSEP obtained with a small number of wavelengths simplify and make more practical the methodology when applied at landscape level.

Table 6-3

Comparison between the predictive regression models at landscape level on Cambisol soil based on the selected wavelength intervals.

Soil fertility parameter	Regression model	R^2 C [a]	R^2 CV [b]	R^2 P [c]	RMSEC [d]	RMSECV [e]	RMSEP [f]	RPD [g]	RER [h]	C Bias [i]	CV Bias [j]	P Bias [k]
	PLS	0.95	0.95	0.96	0.08	0.09	0.09	5.11	19.11	-0.00	-0.00	0.00
OM	**LWR**	**0.98**	**0.98**	**0.97**	**0.06**	**0.07**	**0.08**	**5.80**	**21.67**	**0.01**	**0.01**	**0.01**
	SVM	0.99	0.97	0.96	0.05	0.07	0.11	4.18	15.64	-0.00	-0.00	0.00
	PLS	0.79	0.77	0.81	2.98	3.07	3.30	2.29	7.99	-0.00	0.00	0.26
K_2O	**LWR**	**0.84**	**0.80**	**0.82**	**2.56**	**2.92**	**3.25**	**2.32**	**8.09**	**0.16**	**0.13**	**0.04**
	SVM	0.84	0.80	0.82	2.60	2.92	3.30	2.28	7.97	-0.01	-0.02	0.01
	PLS	0.69	0.68	0.70	0.29	0.31	0.34	1.87	6.39	-0.00	-0.00	-0.05
Olsen P	LWR	0.86	0.76	0.87	0.21	0.26	0.23	2.78	9.49	0.01	0.02	-0.04
	SVM	**0.83**	**0.80**	**0.89**	**0.23**	**0.24**	**0.22**	**3.14**	**10.71**	**-0.02**	**-0.02**	**-0.09**
	PLS	0.77	0.76	0.69	2.47	2.54	3.29	1.73	6.00	-0.00	0.02	0.43
Oniani P	LWR	0.81	0.77	0.68	2.28	2.53	3.38	1.70	5.91	0.28	0.33	0.68
	SVM	**0.79**	**0.76**	**0.71**	**2.36**	**2.53**	**3.04**	**1.86**	**6.45**	**0.01**	**-0.00**	**0.18**

[a] R^2 C- coefficient of determination of calibration; [b] R^2 CV- coefficient of determination of cross validation; [c] R^2 P- coefficient of determination of prediction
[d] RMSEC- root mean square error of calibration; [e] RMSECV- root mean square error of cross validation; [f] RMSEP- root mean square error of prediction;
[g] RPD- ratio of performance deviation; [h] RER- ratio of error range;
[i] C Bias- systematic deviation of calibration; [j] CV Bias- systematic deviation of cross validation; [k] P Bias- systematic deviation of prediction
Data in bold- Best regression model based on the R^2, RMSECV, RMSEP, RPD, RER

6.3.3. Prediction of soil fertility parameters in Vertisol

In Figure 6-4 is shown the spectral reflectance of Vertisol for both, calibration and validation sets. In coincidence with the Cambisol group the final selected pre-processing methods for correcting non linearities, electronic noise and removing variation in the spectra included Log (1/R), smoothing and Mean Centre. In validation set a decreasing peak was observed around the wavelength of 1350 and 1450 nm. This peak is mostly by O-H, C-H and N-H bonds.

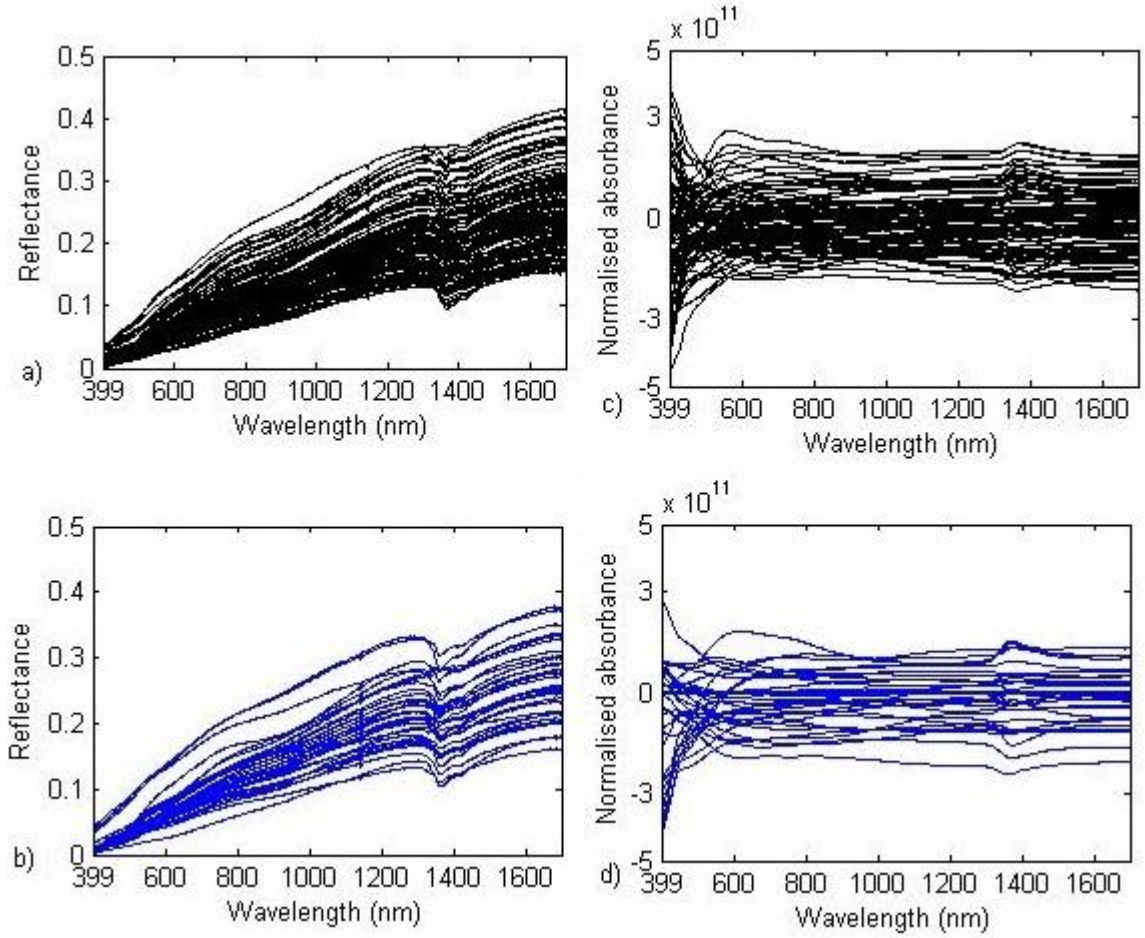


Figure 6-4. Soil spectral reflectance before and after preprocessing for measuring the average fertility at landscape level on Vertisol soil. (a) Raw data of calibration, (b) Preprocessed data of calibration, (c) Raw data of validation (d) Preprocessed data of validation.

In Figure 6-5 the predicted vs. measured plots for OM (SVM) and K_2O (LWR) in Vertisol are respectively illustrated. In agreement with the results obtained for Cambisol in OM the data points were closely grouped in calibration and test sets. In the test set both models (LWR) predicted available Olsen P and Oniani P, with coefficients of determination of 0.80 and 0.81 respectively. In both models the data points were more scattered in both sets than those from OM and K_2O in this soil group.

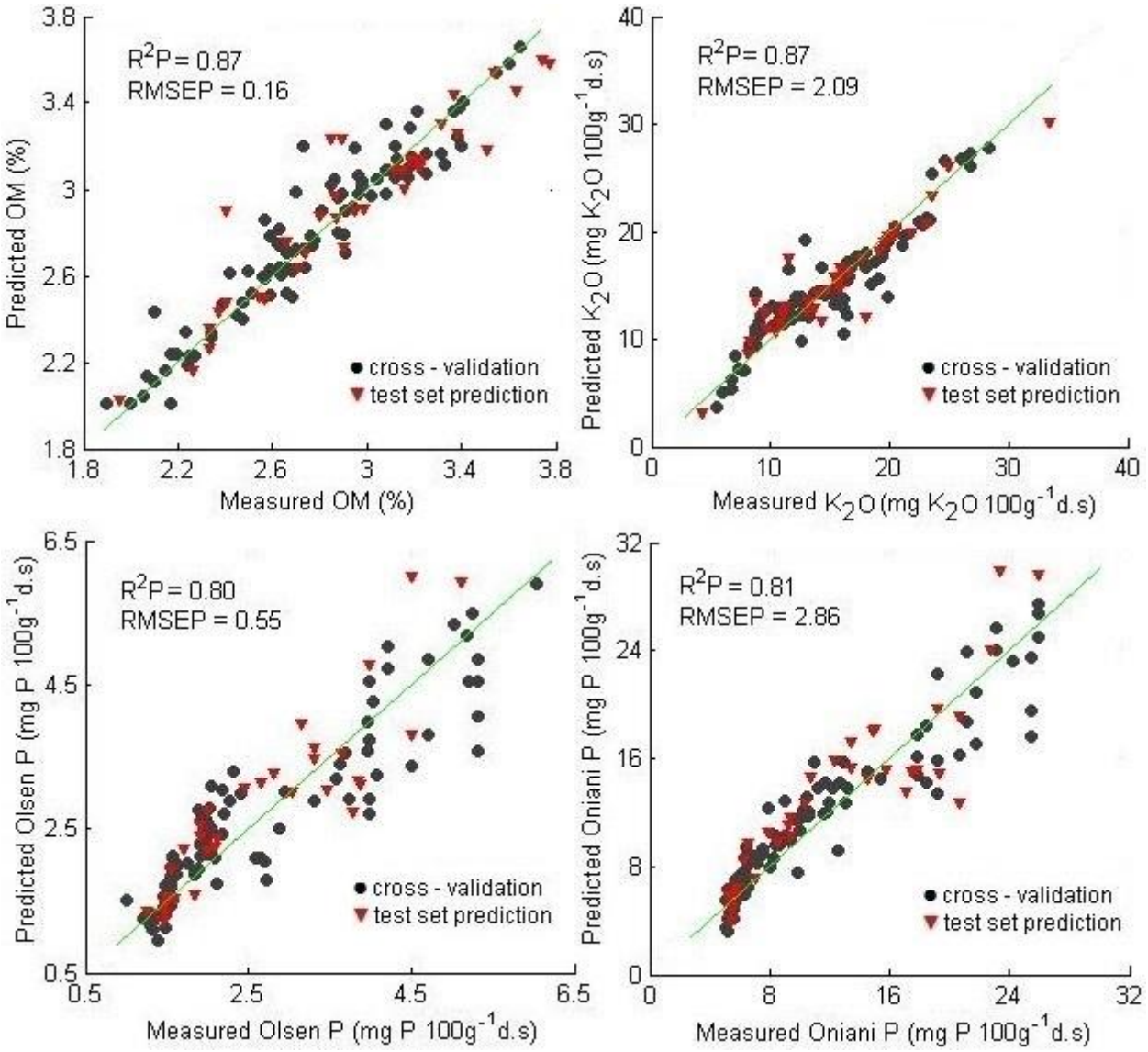


Figure 6-5. Vis/NIR predicted versus chemical conventional analyses of soil fertility parameters at landscape level on Vertisol soil; OM (SVM), K_2O (LWR), Olsen P (LWR), Oniani P (LWR).

The calibration and validation statistics of the best Vis/NIR calibration models obtained with the different calibration methods on the selected Vertisol soil samples are summarised in Table 6-4. The prediction performance of the regression models for OM is lower ($0.81 \leq R^2 \leq 0.87$) than for the Cambisol models. The level of organic matter in soil can influence its reflective properties. Organic matter can also change the texture of the soil, which may cause the soil to reflect differently.

However, the results obtained for K_2O ($0.83 \leq R^2 \leq 0.87$) are better than those obtained for Cambisol. While for Olsen P the prediction performance was comparable to that obtained in Cambisols, the prediction performance for Oniani P was considerably better ($0.74 \leq R^2 \leq 0.81$). In this soil group there was a good agreement among the results for phosphorus, derived from Olsen ($R^2 = 0.80$) and Oniani ($R^2 = 0.81$) methods.

In line with the results obtained for Cambisol for all soil fertility parameters better prediction results were obtained with LWR and SVM (nonlinear regression models) than with PLS (linear regression model).

Also, the high R^2 for potassium and phosphorus obtained for the Vertisol landscape due to the strong correlations between these soil fertility parameters concentration measured by wet chemical analyses and OM (0.82; 0.78; 0.83 for K_2O, Olsen P and Oniani P respetively).

The predictive accuracy of the best regression models was classified as moderately successful for all the soil fertility parameters according to the criteria suggested by Malley et al. (2004).The results are consistent and showed applicable calibrations to predict these standard soil properties based on Vis/NIR spectroscopy.

Table 6-4

Comparison between the best predictive models with others calibrated at landscape level on Vertisol soil

Soil fertility parameter	Regression model	R^2 C [a]	R^2 CV [b]	R^2 P [c]	RMSEC [d]	RMSECV [e]	RMSEP [f]	RPD [g]	RER [h]	C Bias [i]	CV Bias [j]	P Bias [k]
	PLS	0.85	0.84	0.81	0.15	0.16	0.22	2.06	8.31	-0.00	0.00	0.02
OM	LWR	0.89	0.85	0.87	0.13	0.15	0.17	2.66	10.72	-0.00	-0.00	0.01
	SVM	**0.91**	**0.88**	**0.87**	**0.12**	**0.14**	**0.16**	**2.82**	**11.40**	**0.02**	**0.14**	**-0.01**
	PLS	0.85	0.84	0.86	2.01	2.08	2.17	2.71	13.49	0.00	-0.01	0.18
K_2O	**LWR**	**0.87**	**0.85**	**0.87**	**1.93**	**2.07**	**2.09**	**2.80**	**13.96**	**-0.06**	**-0.10**	**0.06**
	SVM	0.87	0.86	0.83	1.91	1.96	2.40	2.46	12.26	-0.00	-0.03	0.32
	PLS	0.78	0.77	0.76	0.59	0.61	0.58	2.20	8.66	0	0.00	0.04
Olsen P	**LWR**	**0.84**	**0.81**	**0.80**	**0.51**	**0.56**	**0.55**	**2.42**	**9.52**	**0.04**	**0.05**	**0.16**
	SVM	0.90	0.79	0.76	0.40	0.60	0.67	1.93	7.60	-0.05	-0.02	0.12
	PLS	0.84	0.83	0.77	2.59	2.67	3.10	2.00	6.68	0	-0.01	-0.01
Oniani P	**LWR**	**0.89**	**0.86**	**0.81**	**2.14**	**2.40**	**2.86**	**2.24**	**7.48**	**0.26**	**0.23**	**0.72**
	SVM	0.93	0.88	0.74	1.07	2.27	3.10	2.00	6.68	-0.10	-0.35	0.03

[a] R^2 C- coefficient of determination of calibration; [b] R^2 CV- coefficient of determination of cross validation; [c] R^2 P- coefficient of determination of prediction
[d] RMSEC- root mean square error of calibration; [e] RMSECV- root mean square error of cross validation; [f] RMSEP- root mean square error of prediction;
[g] RPD- ratio of performance deviation; [h] RER- ratio of error range;
[i] C Bias- systematic deviation of calibration; [j] CV Bias- systematic deviation of cross validation; [k] P Bias- systematic deviation of prediction
Data in bold- Best regression model based on the R^2, RMSECV, RMSEP, RPD, RER

6.3.4. Identification of effective wavelength intervals in Vertisol

Vis/NIR soil reflectance measurements over a wide spectral range can provide the information needed to quantify different soil fertility parameters. However, the selection of some wavelength ranges represents a more efficient approach to analyse the parameters of interest. Therefore, the number of wavelengths was reduced to 20% of the full spectrum (84 wavelengths). Therefore, three wavelength intervals were selected as important for each soil fertility parameter (Figure 6-6).

In general, different wavelengths were found to be important for the different soil fertility parameters. There were some wavelengths in common or similar among these parameters. The interval number 1 (399 – 489 nm) was selected for OM, Olsen P and Oniani P. Also, the interval number 2 (489 – 580 nm) was common to all the soil fertility parameters in Vertisol landscape. All these intervals fit in the Vis region (380 – 780 nm).

Wavelengths included in the intervals number 1 and 2 selected for OM are in line with the results reported by several authors. In Tokat Province from northern Turkey Volkan et al. (2010) obtained the best predictive models for organic matter (R^2 =0.80) within the visible range. Kirshnan et al. (1980) as well as Cozzolino and Morón (2006) considered the VIS range important for calibration of OM. Good estimations between the OM content and spectral reflectance were obtained by Lee et al. (2003) at 428 nm and 512 nm in soils from Florida. In China Bocheng (2004) indicated that the OM content was significantly correlated to the original reflectance, logarithm for reciprocal reflectance in 447 nm, and the first differentiation of logarithm for reciprocal reflectance in the band of 516 nm. According to Yu et al. (2013) the wavelength of 492 nm possessed the best prediction accuracy reported for OM content. This wavelength is included in the interval number 2 selected in this study.

Moreover, wavelengths included in the intervals selected for K_2O, Olsen P and Oniani P have been reported by other researchers. In two agricultural fields in north-eastern Mississippi, Thomasson et al. (2001) selected the wavelengths 525 nm, 575 nm, 625 nm and 675 nm for potassium (K). Also, for phosphorus (P) these authors selected the wavelengths 475 nm, 525 nm, 575 nm, 625 nm, 1075 nm, 1125 nm. For K, Lee et al. (2003) reported the wavelengths 522 nm, 690 nm and 740 nm. While for P the same authors reported the wavelengths 428 nm, 430 nm, 522 nm and 1100 nm.

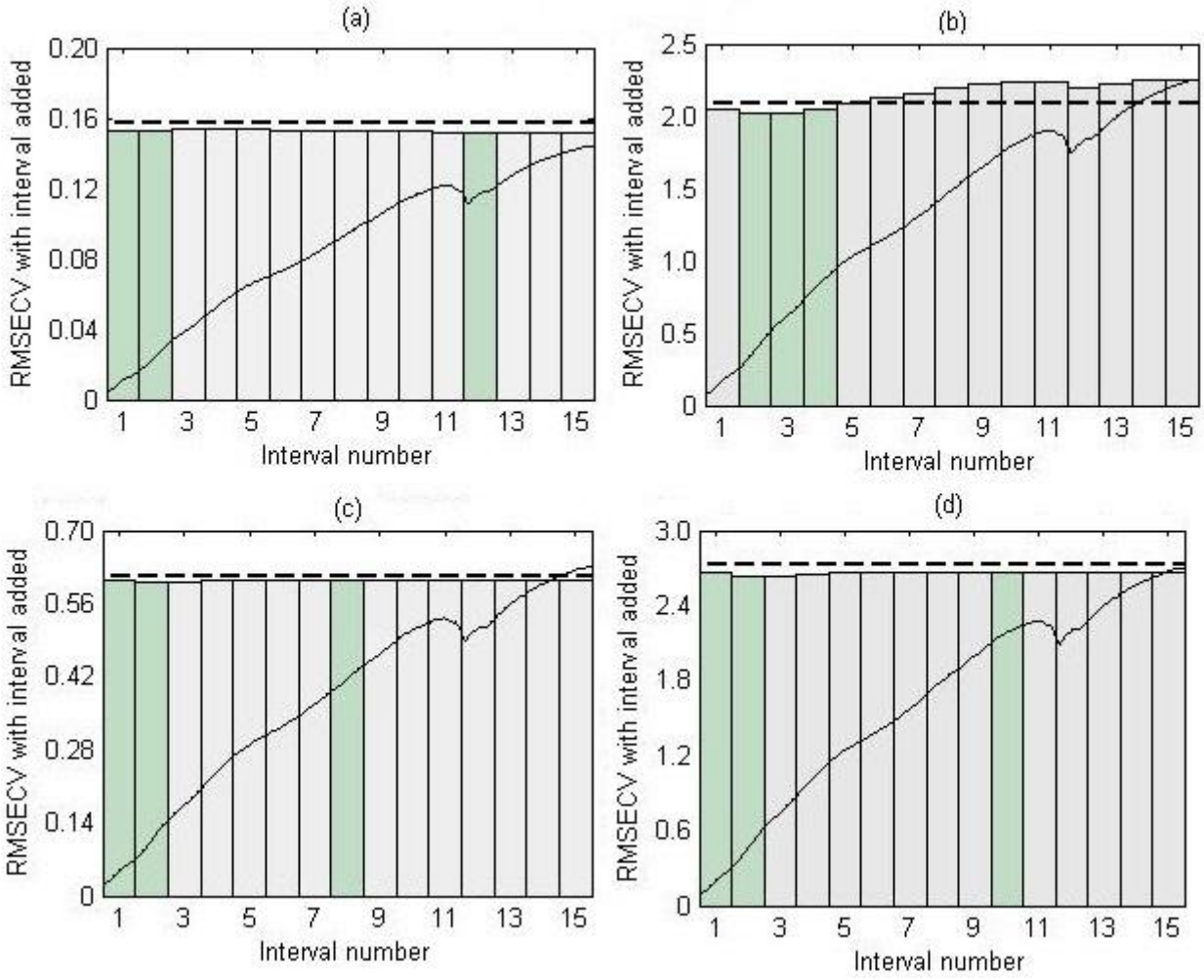


Figure 6-6. Spectral region selected (dark colour) by Forward Interval PLS model (iPLS) at landscape level on Vertisol soil.

Intervals selected:

(a) OM- [1 (399 – 489 nm), 2 (489 – 580 nm), 12 (1402 – 1480 nm)]
(b) K_2O- [2 (489 – 580 nm), 3 (580 – 674 nm), 4 (674 – 764 nm)]
(c) Olsen P- [1 (399 – 489 nm), 2 (489 – 580 nm), 8 (1050 – 1145 nm)]
(d) Oniani P- [1 (399 – 489 nm), 2 (489 – 580 nm), 10 (1235 – 1320 nm)]

Table 6-5 compares the accuracy of the predictive regression models at landscape level using selected wavelength intervals. The three regression models (PLS, SVM, LWR) developed for OM achieve a moderately successful prediction performance. Also, SVM and LWR models gave the more accurate results with the highest R^2 of prediction (0.88) and the lowest RMSEP (0.15). The other soil fertility parameters also increased the R^2 of prediction, RPD and kept the model selected as the best with the full spectrum (LWR). Also, neither of the soil fertility parameters differ as to the category related to the accuracy of performance observed with the full spectrum. The results demonstrated that all models developed for the soil fertility parameters using just 20% of the wavelengths performed better than with the full spectrum. Thus, the highest accuracy levels and the lowest RMSEP obtained with a small number of wavelengths simplify and make more practical the methodology when applied at landscape level. Therefore, at landscape level on Vertisol the wavelength selection exhibited an advantage over the full spectrum in terms of accuracy of the models.

Table 6-5

Comparison between the predictive regression models at landscape level on Vertisol soil based on the selected wavelength intervals

Soil fertility parameter	Regression model	R^2 C [a]	R^2 CV [b]	R^2 P [c]	RMSEC [d]	RMSECV [e]	RMSEP [f]	RPD [g]	RER [h]	C Bias [i]	CV Bias [j]	P Bias [k]
	PLS	0.86	0.84	0.83	0.16	0.16	0.21	2.17	8.76	0.00	-0.00	0.03
OM	LWR	0.89	0.87	0.88	0.14	0.15	0.15	3.00	12.13	-0.00	-0.01	0.00
	SVM	**0.89**	**0.88**	**0.88**	**0.14**	**0.15**	**0.15**	**3.01**	**12.16**	**0.00**	**0.00**	**0.01**
	PLS	0.85	0.84	0.85	2.01	2.05	2.22	2.64	13.15	0.00	0.00	-0.09
K_2O	**LWR**	**0.89**	**0.85**	**0.88**	**1.73**	**2.00**	**2.03**	**2.89**	**14.38**	**-0.13**	**-0.15**	**-0.09**
	SVM	0.87	0.87	0.85	1.85	1.89	2.29	2.56	12.74	-0.06	-0.03	-0.01
	PLS	0.78	0.77	0.78	0.57	0.59	0.55	2.33	9.18	0.00	-0.00	0.07
Olsen P	**LWR**	**0.87**	**0.81**	**0.83**	**0.45**	**0.54**	**0.50**	**2.75**	**10.83**	**0.05**	**0.07**	**0.19**
	SVM	0.90	0.81	0.77	0.38	0.54	0.56	2.28	9.00	-0.04	-0.05	0.06
	PLS	0.85	0.84	0.80	2.55	2.63	2.87	2.16	7.22	0	-0.00	0.09
Oniani P	**LWR**	**0.92**	**0.88**	**0.82**	**1.87**	**2.29**	**2.75**	**2.35**	**7.87**	**0.26**	**0.35**	**0.80**
	SVM	0.95	0.90	0.79	1.52	2.22	2.96	2.14	7.14	-0.18	-0.20	0.59

[a] R^2 C- coefficient of determination of calibration; [b] R^2 CV- coefficient of determination of cross validation; [c] R^2 P- coefficient of determination of prediction
[d] RMSEC- root mean square error of calibration; [e] RMSECV- root mean square error of cross validation; [f] RMSEP- root mean square error of prediction;
[g] RPD- ratio of performance deviation; [h] RER- ratio of error range;
[i] C Bias- systematic deviation of calibration; [j] CV Bias- systematic deviation of cross validation; [k] P Bias- systematic deviation of prediction
Data in bold- Best regression model based on the R^2, RMSECV, RMSEP, RPD, RER

6.4. Conclusions

In this chapter, the potential of Vis/NIR spectroscopy for prediction the average soil fertility parameters of Cuban agricultural fields with Cambisol or Vertisol soil has been evaluated. One preprocessing method had been applied for this purpose. Althought the results showed that Vis/NIR spectroscopy could be useful for rapidly determining the OM (%); Olsen P; Oniani P and K_2O (mg 100 g^{-1} d.s) content in these Cuban agricultural soils, future measurements should benefit from lessons learned in this initial attempt. Those lessons are related to the fact that the soil fertility parameters studied could be better estimated by using non-linear regression models with a successful (OM; R^2 of 0.93 with LWR), moderately successful (Olsen P; R^2 of 0.81 whit SVM), moderately useful (K_2O; R^2 of 0.79 with LWR) and less reliable (Oniani P; R^2 of 0.65 with SVM) accuracy for the Cambisol soil samples. For Vertisol soil samples the non-linear regression models were moderately successful for all considered soil fertility parameters, with R^2 of 0.87; 0.87; 0.80 and 0.81 for OM (SVM), K_2O (LWR), Olsen P (LWR) and Oniani P (LWR) respectively. The soil fertility parameters K_2O, Olsen P and Oniani P are not spectrally active in the Vis/NIR range, and then they were most likely predicted through their strong correlation with the OM. The OM content can be predicted from the Vis/NIR spectra thanks to the spectral activity of the CH-bonds.

This research has also shown that the PLS method consistently identified spectral regions of interest for better estimating each soil fertility parameter. The advantage of the selection of the wavelength intervals is that the full spectrum need not be used. Thus, the number of wavelengths was reduced to 84, which is the 20% of the full spectrum. In general, different wavelengths were found to be important for the different soil fertility parameters. There were some wavelengths in common among K_2O and Olsen P (399 – 674 nm) and among all soil fertility parameters (399 – 489 nm) in Cambisol soil. Also, in Vertisol soil there were wavelength intervals in common among all soil fertility parameters (489 – 580 nm) and among OM, Olsen P and Oniani P (399 – 489 nm). As a general recommendation from this study, the soil variation and the spatial variability across this landscape area must be taken into consideration for obtaining successful, useful or reliable Vis/NIR prediction models of these soil fertility parameters in air dried samples.

Chapter 7

Prediction of variation in soil fertility parameters within a field

7.1. Introduction

Since productivity is influenced by the soil fertility parameters, its variation within a field can lead to variations in the productivity. By assessing this variability the management of the soil can be optimized in a more effective way (Mzuku et al., 2005).

Precision agriculture, or site-specific crop management (SSCM), is an information-based management intensive approach to farming. Instead of managing a field as a whole, the philosophy of precision agriculture is to manage individual areas within a field differently. Accounting for soil variability is a critical need for SSCM (Hong et al., 2002).

Vis/NIR spectroscopy has been suggested to be an efficient tool to predict, within a field, soil fertility parameters that can be of significant value when establishing agricultural field trials and in precision farming (Wang et al., 2008). Accordingly, Bricklemyer and Brown (2010) reported that lab-based Vis/NIR spectroscopy is a viable technique for estimating various soil properties.

This indicates that Vis/NIR spectroscopy could play an important role in optimizing the fertilisation for crop production. Specifically in sugarcane cultivation, where a balanced mineral nutrition is very important in order to achieve high yields with quality, this technology could provide a large added value.

In this chapter the potential of Vis/NIR spectroscopy for predicting the variation in soil fertility parameters within a field is evaluated for two agricultural fields, one with a Cambisol soil and one with a Vertisol soil, in the Villa Clara province of Cuba. The materials and methods are illustrated in section 7.2, while the results and discussion are presented in section 7.3. Finally, conclusions are drawn from this research in section 7.4.

7.2. Materials and methods

The 189 Cambisol soil samples from 189 different agricultural fields, which were also used in Chapter 6, were used to build calibration models between the Vis/NIR spectra and the constituent concentrations. The test set consisted of 37 Cambisol soil samples sampled separately, randomly and independently from different locations distributed over the one field. Similarly, the 144 Vertisol soil samples from 144 different fields, also used in Chapter 6, were used to build calibration models. These models were then tested for predicting the soil fertility parameters of 29 Vertisol soil samples sampled from one field. The whole sampling process for acquiring these soil samples has been described in Chapter 3, while the spectral measurements were taken from air-dried soil samples in the Vis/NIR region (399 – 1697 nm) with a diode array spectrophotometer (CORONA PLUS REMOTE Vis/NIR SB, Zeiss, Jena, Germany) as described in Chapter 4. Similar to the procedure followed in Chapter 6, the optimal data preprocessing method in terms of the robustness and accuracy of multivariate analysis was selected. This selection was based on the differences in accuracy among different pre-processing steps. Therefore, the prediction performance of the calibration models was assessed based on the R^2 and RMSE calculated in cross-validation at the landscape level and test set prediction at the field level. Calibration accuracy was again assessed using the guidelines proposed by Malley et al. (2004) and Nduwamungu et al. (2009), which were presented in Chapter 6. Also, PLS, SVM and LWR were tested as calibration techniques in Venetian blinds cross-validation. All calculations were performed with the PLS toolbox (Eigenvector inc., Wenatchee, WA) in MATLAB 7.9 (R2009b, The Mathworks, Nattick, MT).

Some outliers in the calibration set were identified based on their large values on the scatter plot. According to Esbensen et al. (2002) those samples which lie far away from the others or which do not fit in well with the rest in the plot, are possible outliers, because they are different from the others. In this research, a possible explanation for these outliers could be related to the sampling procedure or chemical measurement errors. Even if every precaution is taken to avoid systematic errors, small deviations or random errors arise that are unavoidable and not identifiable. Random errors, by definition, are impossible to illustrate. As these outliers, might have a detrimental effect on the calibration models they were removed from the dataset. After removing these samples from the calibration set, the regression models were recalculated.

7.3. Results and discussion

7.3.1. Prediction of soil fertility parameters within a field on Cambisol soil

The raw and preprocessed data for the different Cambisol samples taken from one field are illustrated in Figure 7-1. In line with the results obtained at landscape level the final selected pre-processing method was Log (1/R), Smoothing and Mean Centre. The reflectance spectra at landscape level (a) and within a field (b) have the same general characteristics, which suggest that the soil type could be derived from the spectra.

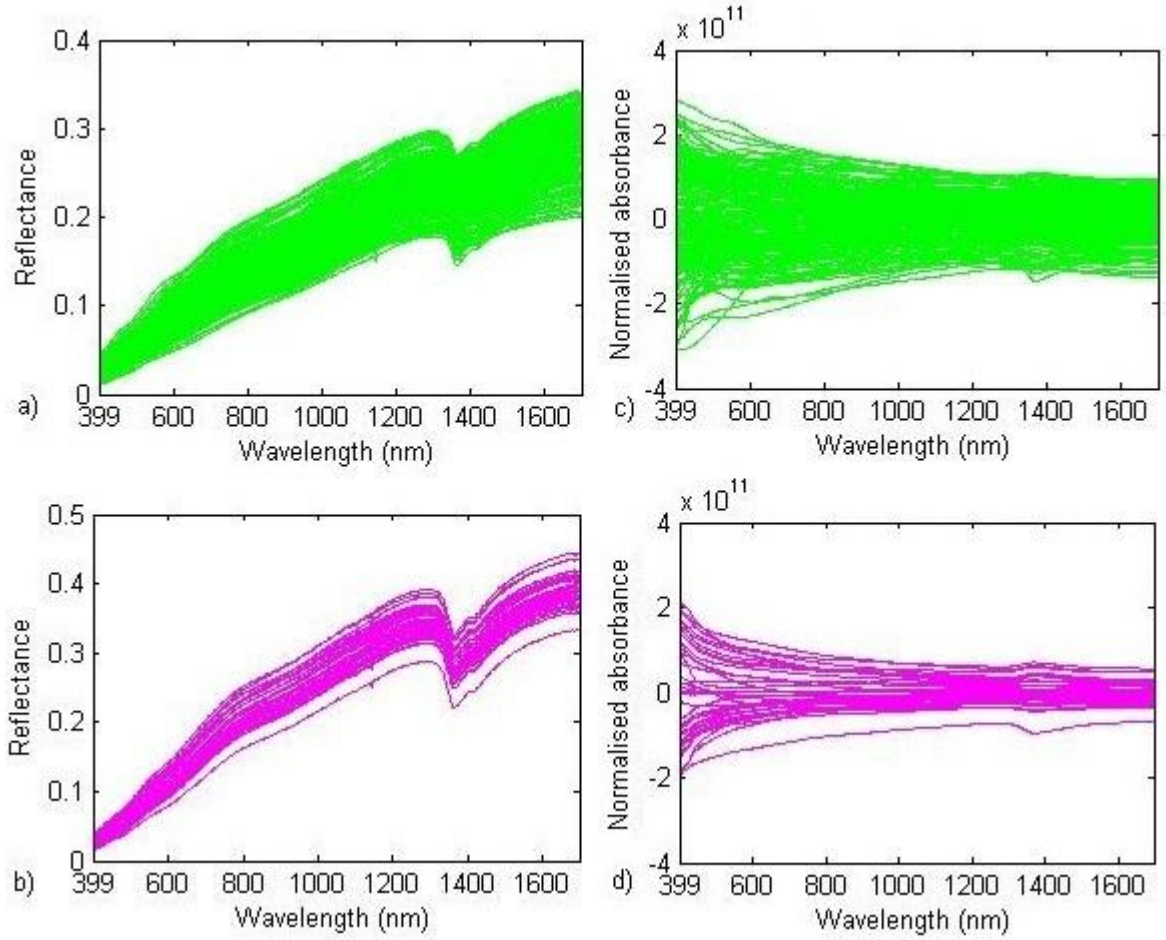


Figure 7-1. Soil spectral reflectance before and after preprocessing for predicting the variation in soil fertility parameters within a field on Cambisol soil. (a) Raw data of calibration, (b) Preprocessed data of calibration, (c) Raw data of validation, (d) Preprocessed data of validation.

In Figure 7-2 the scatter plots of the better predicted vs. measured values of all the soil fertility parameters within a field are shown. Across the full range, OM content (SVM, R^2 of 0.92) exhibited a good agreement between calibration and validation data at lower and medium values than at higher values of the validation set. On the other hand for K_2O (LWR, R^2 of 0.63) the data points tend to be scattered along the whole range of the model. The best regression models obtained for Olsen P (SVM, R^2 of 0.83) and Oniani P (LWR, R^2 of 0.72) illustrated more accurate predictions for these parameters. These higher predictions could also be explained by the relatively higher correlations of these soil fertility parameters with OM (r = 0.85 and 0.83 for Olsen P and Oniani P, respectively).

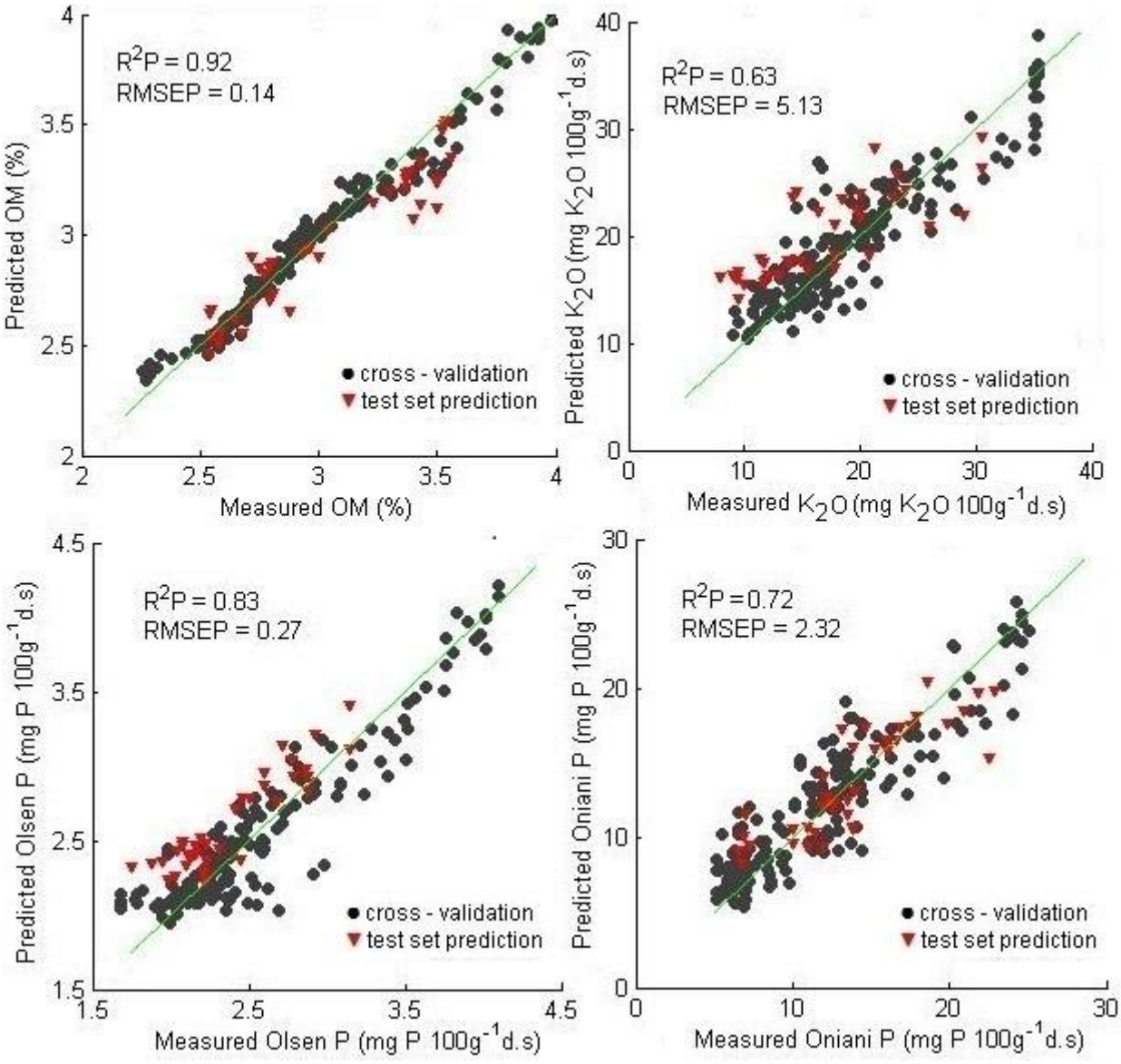


Figure 7-2. Vis/NIR predicted versus chemical conventional analyses of soil fertility parameters within a field on Cambisol soil; OM (SVM), K_2O (LWR), Olsen P (SVM), Oniani P (LWR).

By comparison, the nonlinear regression models produce better predictions within a field as well as at landscape level. It is worth noting that SVM model for OM achieves similar accuracy within a field (R^2 of 0.92; RMSEP of 0.14) than LWR at landscape level (R^2 of 0.93; RMSEP of 0.12). LWR model for K_2O shows a lower prediction within a field (R^2 of 0.63; RMSEP of 5.13) than at landscape level (R^2=0.79; RMSEP=3.47).

Also, the prediction of the best models (SVM) for Olsen P produces moderately successful performance in both scenarios (R^2 of 0.83; RMSEP of 0.27 within a field and R^2 of 0.83; RMSEP of 0.27 at landscape level). Finally, the numerical values of statistics for Oniani P obtained with SVM model within a field (R^2 of 0.72; RMSEP of 2.32) are better than those obtained with LWR model at landscape level (R^2 of 0.65; RMSEP of 3.31). Further, when the best models are compared the results demonstrate that for this soil type a calibration made within a field is capable of being used over the landscape scale.

The statistics on the prediction performance of Vis/NIR spectroscopy using PLS, LWR and SVM for the Cambisol field are summarised in Table 9. The best test set prediction results were obtained for organic matter (R^2 of 0.92), while the results for K_2O were worst ($0.61 \leq R^2 \leq 0.63$). The obtained prediction R^2 were $0.68 \leq R^2 \leq 0.83$ for Olsen P and $0.70 \leq R^2 \leq 0.72$ for Oniani P. In general LWR was found to give the best prediction results.

In line with Volkan et al. (2010) the values of R^2 for soil fertility parameters prediction in the validation set were lower and RMSEP values higher than corresponding values in the calibration set. Except for Olsen P the R^2 values were slightly worse in test set prediction than in cross-validation. This could have been expected as the variation within a field is smaller than the variation between different fields at landscape level and some of the unspecific correlations which exist between the spectra and chemical constituents of different fields may not exist within a field (Dardenne et al., 2000; Volkan et al., 2010).

According to the suggested guidelines, there was a coincidence in the predictive accuracy of the best regression models for OM and Olsen P; both were classified as moderately successful. For K_2O the model was less reliable, which could be explained by the lack spectral activity of this parameter. Finally, the prediction of Oniani P was classified as moderately useful.

Table 7-1

Comparison between the best predictive models with others calibrated within a field on Cambisol soil

Soil fertility parameter	Regression model	R^2 C [a]	R^2 CV [b]	R^2 P [c]	RMSEC [d]	RMSECV [e]	RMSEP [f]	RPD [g]	RER [h]	C Bias [i]	CV Bias [j]	P Bias [k]
	PLS	0.96	0.95	0.92	0.08	0.09	0.14	3.22	8.88	-0.00	-0.00	-0.08
OM	LWR	0.96	0.95	0.92	0.07	0.09	0.14	3.45	9.51	0.01	0.02	-0.09
	SVM	**0.97**	**0.97**	**0.92**	**0.07**	**0.07**	**0.14**	**3.45**	**9.51**	**-0.00**	**-0.00**	**-0.09**
	PLS	0.74	0.73	0.61	3.53	3.63	5.88	1.56	5.66	-1.07	0.02	4.30
K_2O	**LWR**	**0.83**	**0.76**	**0.63**	**2.90**	**3.43**	**5.13**	**1.61**	**5.87**	**0.33**	**0.43**	**3.37**
	SVM	0.80	0.75	0.63	3.08	3.53	5.43	1.52	5.53	-0.18	-0.07	3.56
	PLS	0.72	0.70	0.68	0.30	0.31	0.34	1.29	4.97	0	-0.00	0.19
Olsen P	LWR	0.89	0.82	0.77	0.19	0.25	0.34	1.46	5.59	0.02	0.02	-0.23
	SVM	**0.87**	**0.83**	**0.83**	**0.20**	**0.23**	**0.27**	**2.95**	**11.32**	**-0.00**	**-0.00**	**0.24**
	PLS	0.74	0.73	0.70	2.71	2.77	2.52	1.76	6.50	-1.07	0.01	0.26
Oniani P	**LWR**	**0.82**	**0.77**	**0.72**	**2.26**	**2.58**	**2.32**	**1.90**	**7.02**	**0.33**	**0.42**	**-0.06**
	SVM	0.82	0.79	0.71	2.44	2.36	2.48	1.78	6.58	-0.06	-0.01	0.16

[a] R^2 C- coefficient of determination of calibration; [b] R^2 CV- coefficient of determination of cross validation; [c] R^2 P- coefficient of determination of prediction
[d] RMSEC- root mean square error of calibration; [e] RMSECV- root mean square error of cross validation; [f] RMSEP- root mean square error of prediction;
[g] RPD- ratio of performance deviation; [h] RER- ratio of error range;
[i] C Bias- systematic deviation of calibration; [j] CV Bias- systematic deviation of cross validation; [k] P Bias- systematic deviation of prediction
Data in bold- Best regression model based on the R^2, RMSECV, RMSEP, RPD, RER

Bogrekci and Lee (2007) also reported good prediction potential of the P concentration from Vis/NIR reflectance spectra, with R^2 values of 0.93, 0.95 and 0.76 for total, Mehlich-1 and water-soluble P, respectively. In this research the very well predicted P content (Olsen and Oniani) by using Vis/NIR spectroscopy might be related to the higher correlation with OM.

On the other hand, the results for OM and K in this research were very similar to those obtained by He et al. (2007). They achieved good predictions with PLS for OM (R^2 of 0.93). However, they concluded that Vis/NIR spectroscopy was not a good tool for P and K prediction with R^2 values of 0.47 and 0.68, respectively. Also, in the study carried out by Wetterlind et al. (2008) the validation statistics indicated that the Vis/NIR calibrations for OM were reliable (R^2 of 0.89 and R^2 of 0.87).

7.3.2. Identification of effective wavelength intervals within a field on Cambisol soil

Figure 7-3 demonstrates that within a field, as well as at landscape level, the wavelengths included in interval number 1 (399 – 489 nm) were selected for all fertility parameters. Also, there were coincidences as to the three intervals selected for K_2O- number 1; 2 (489 – 580 nm); 3 (580 – 674 nm) and in two of the intervals selected for Oniani P, number 1 and number 8 (1050 – 1145 nm). Therefore, some wavelengths selected within a field and at landscape level on Cambisol soil fall in comparable spectral regions.

Nowkandeh et al. (2013) found 477 nm, 905 nm, 972 nm, 1013 nm, 1023 nm and 1033 nm as most important wavelengths for detecting OM. The first wavelength is included in the interval number 1 and it is in agreement with those observed within a field. The second one belongs to the interval number 6 (860 – 952 nm) and the rest to the interval number 7 (952 – 1050 nm). For the other soil fertility parameters, some of the wavelengths selected by Thomasson et al. (2001) and Lee et al. (2003) are in agreement in both scenarios too. For K_2O there is a total coincidence at landscape level and within a field. For P, there was a coincidence with the wavelengths 428 nm, 430 nm, 475 nm, 1075 nm, 1100 nm and 1125 nm. However, these authors also reported other wavelengths for P (1225 nm, 1412 nm, 1425 nm, 1475 nm, 1498 nm, 1525 nm) which are in agreement with those observed in this study within a field but differ from the wavelengths selected at landscape level. Moreover, the regions included in the intervals selected, are known to correspond to spectral features related to these soil fertility parameters.

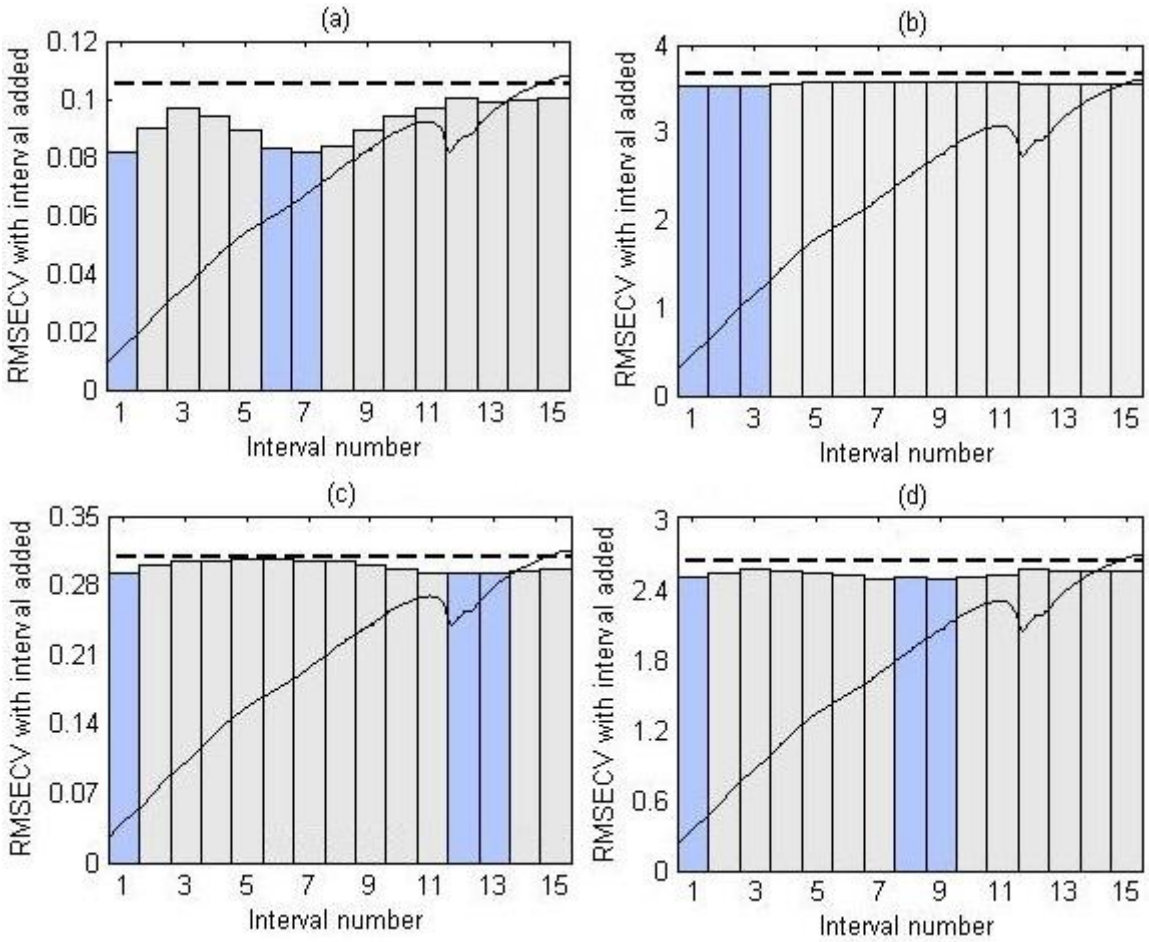


Figure 7-3. Spectral region selected (dark colour) by Forward Interval PLS model (iPLS) within a field on a Cambisol soil.

Intervals selected:

(a) OM- [1 (399 – 489 nm), 6 (860 – 952 nm), 7 (952 – 1050 nm)]
(b) K_2O- [1 (399 – 489 nm), 2 (489 – 580 nm), 3 (580 – 674 nm)]
(c) Olsen P- [1 (399 – 489 nm), 12 (1402 – 1480 nm), 13 (1480 – 1555 nm)]
(d) Oniani P- [1 (399 – 489 nm), 8 (1050 – 1145 nm), 9 (1145 – 1235 nm)]

In table 7-2 the results of the predictive regression models within a field based on the selected wavelength intervals are shown. The three regression models (PLS, SVM, LWR) developed for OM achieve a successful prediction performance. It was the same result obtained with the full spectrum within a field. However, the RMSEP values were increased as a consequence of reducing the number of wavelengths. Both results differ from those obtained at landscape level, where the three regression models improve the prediction accuracy and reduce the RMSEP values.

By reducing the number of wavelengths, the soil fertility parameters K_2O and Olsen P increased the R^2 of prediction, reduced the RMSEP and kept the model selected as the best with the full spectrum (LWR – K_2O; SVM – Olsen P). Also, the prediction accuracy was the same as observed with the full spectrum, except for LWR (R^2 of 0.82 – moderately successful). For Oniani P the R^2 of prediction and RMSEP values were increased. The best model was SVM and not LWR as obtained with the full spectrum.

The RMSEP values were increased only within a field and just for OM and Oniani P. In this sense Figure 7-3 shows that the wavelength intervals identified were not identical for OM (interval number 6) and Oniani P (interval number 8). It means that a higher RMSECV value than in the others was obtained in this respective interval. It is probable that using data only from the two intervals with a lower RMSECV would give better results of RMSEP for both soil fertility parameters. This verification would be an important additional step toward the improvement of the methodology which pursues the use of a low number of wavelengths for obtaining better or similar prediction accuracy.

Table 7-2

Comparison between the predictive regression models within a field on Cambisol soil based on the selected wavelength intervals

Soil fertility parameter	Regression model	R^2 C	R^2 CV	R^2 P	RMSEC [d]	RMSECV [e]	RMSEP [f]	RPD [g]	RER [h]	C Bias [i]	CV Bias [j]	P Bias [k]
	PLS	0.95	0.95	0.92	0.08	0.08	0.15	3.31	9.12	-0.00	-0.00	-0.10
OM	LWR	0.98	0.97	0.92	0.05	0.07	0.18	3.72	10.25	0.01	0.01	-0.15
	SVM	**0.98**	**0.96**	**0.92**	**0.05**	**0.06**	**0.18**	**3.72**	**10.25**	**-0.00**	**0.07**	**-0.14**
	PLS	0.76	0.73	0.63	3.36	3.53	5.50	1.56	5.69	-0.00	-0.00	3.79
K_2O	**LWR**	**0.83**	**0.75**	**0.65**	**2.79**	**3.47**	**4.47**	**2.88**	**10.47**	**0.17**	**0.30**	**3.91**
	SVM	0.80	0.76	0.64	3.04	3.33	4.53	1.65	5.99	-0.35	-0.19	2.48
	PLS	0.74	0.73	0.69	0.29	0.29	0.30	1.29	4.95	-0.00	0.00	0.10
Olsen P	LWR	0.87	0.83	0.82	0.21	0.24	0.25	1.63	6.24	0.02	0.03	-0.11
	SVM	**0.86**	**0.82**	**0.84**	**0.21**	**0.24**	**0.20**	**2.40**	**9.21**	**0.01**	**-0.00**	**0.13**
	PLS	0.79	0.78	0.73	2.42	2.50	2.71	1.71	6.34	-0.00	-0.01	-0.86
Oniani P	LWR	0.84	0.79	0.73	2.12	2.44	2.69	2.08	7.70	0.19	0.19	-1.66
	SVM	**0.83**	**0.81**	**0.72**	**2.20**	**2.34**	**2.54**	**1.90**	**7.02**	**-0.08**	**-0.06**	**-1.03**

[a] R^2 C- coefficient of determination of calibration; [b] R^2 CV- coefficient of determination of cross validation; [c] R^2 P- coefficient of determination of prediction
[d] RMSEC- root mean square error of calibration; [e] RMSECV- root mean square error of cross validation; [f] RMSEP- root mean square error of prediction;
[g] RPD- ratio of performance deviation; [h] RER- ratio of error range;
[i] C Bias- systematic deviation of calibration; [j] CV Bias- systematic deviation of cross validation; [k] P Bias- systematic deviation of prediction
Data in bold- Best regression model based on the R^2, RMSECV, RMSEP, RPD, RER

7.3.3. Prediction of soil fertility parameters within a field on Vertisol soil

The raw and preprocessed data for the prediction of soil fertility parameters within a field on Vertisol soil are shown in Figure 7-4. The pre-processing method included Log (1/R), Smoothing and Mean Centre. Diffuse reflectance in the Vis/NIR region might be influenced by physical properties such as the size and arrangement of particles. This clayey soil has water molecules in its chemical composition and not all water molecules are removed by the process of air drying. Therefore, even dry, this clayey soil dips in water absorption bands around the 1400 nm, which is more significant within a field.

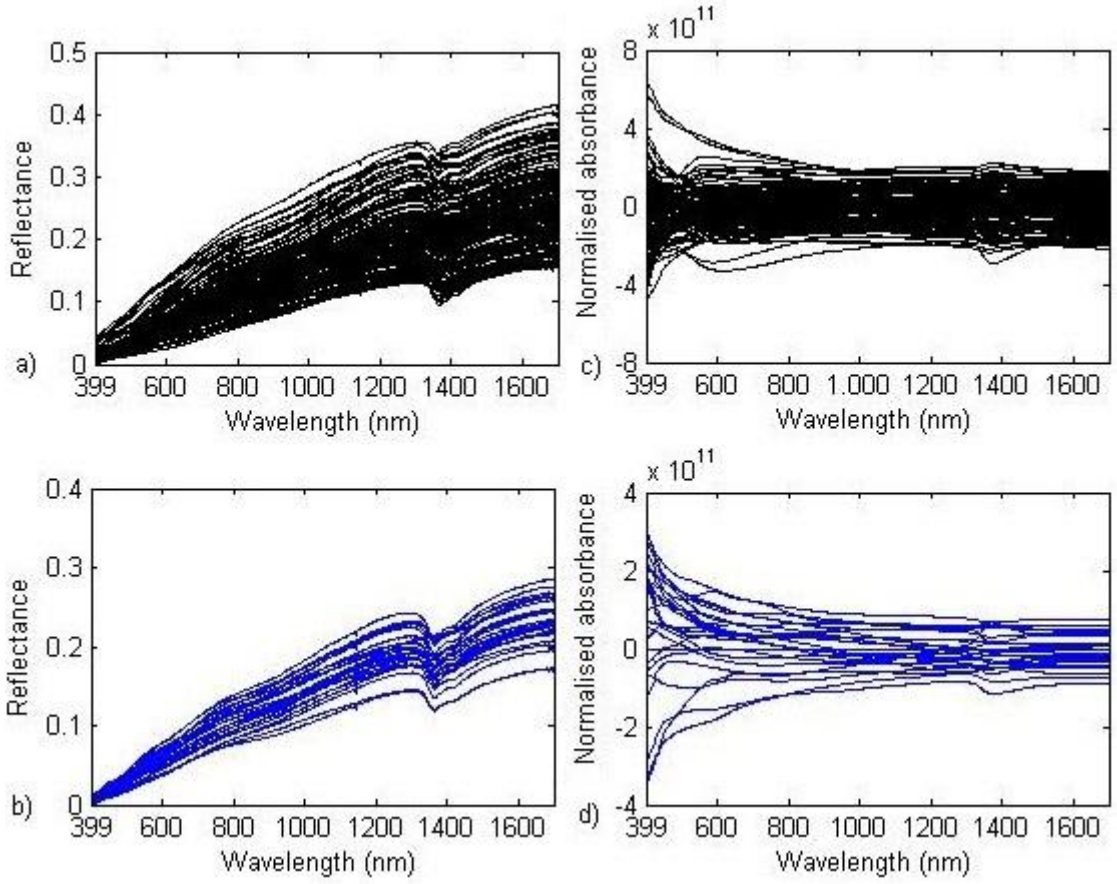


Figure 7-4. Soil spectral reflectance before and after preprocessing for predicting the variation in soil fertility parameters within a field on Vertisol soil. (a) Raw data of calibration, (b) Preprocessed data of calibration, (c) Raw data of validation, (d) Preprocessed data of validation.

The SVM regression model obtained the best results for OM and K_2O with the higher coefficient of determination (R^2= 0.79) for OM (Figure 7-5). The Olsen P and Oniani P methods applied for the determination of phosphorus gave the same R^2 value of 0.58. Except in K_2O, the prediction accuracy in the rest of soil fertility parameters were considerable lower than the observed in Cambisol. A possible explanation for this might be the soil moisture content holded in the soil samples. In high clay soils like Vertisols, the air spaces are numerous and smaller, making it hard for water to drain out. Therefore, high clay content soil has a high water-holding capacity (even dry). Thus, soil moisture can affect the reflectance of the soil in a variable way across the electromagnetic spectrum. Soil reflectance can decrease while soil moisture content increase, or increases when the soil moisture reaches a certain content.

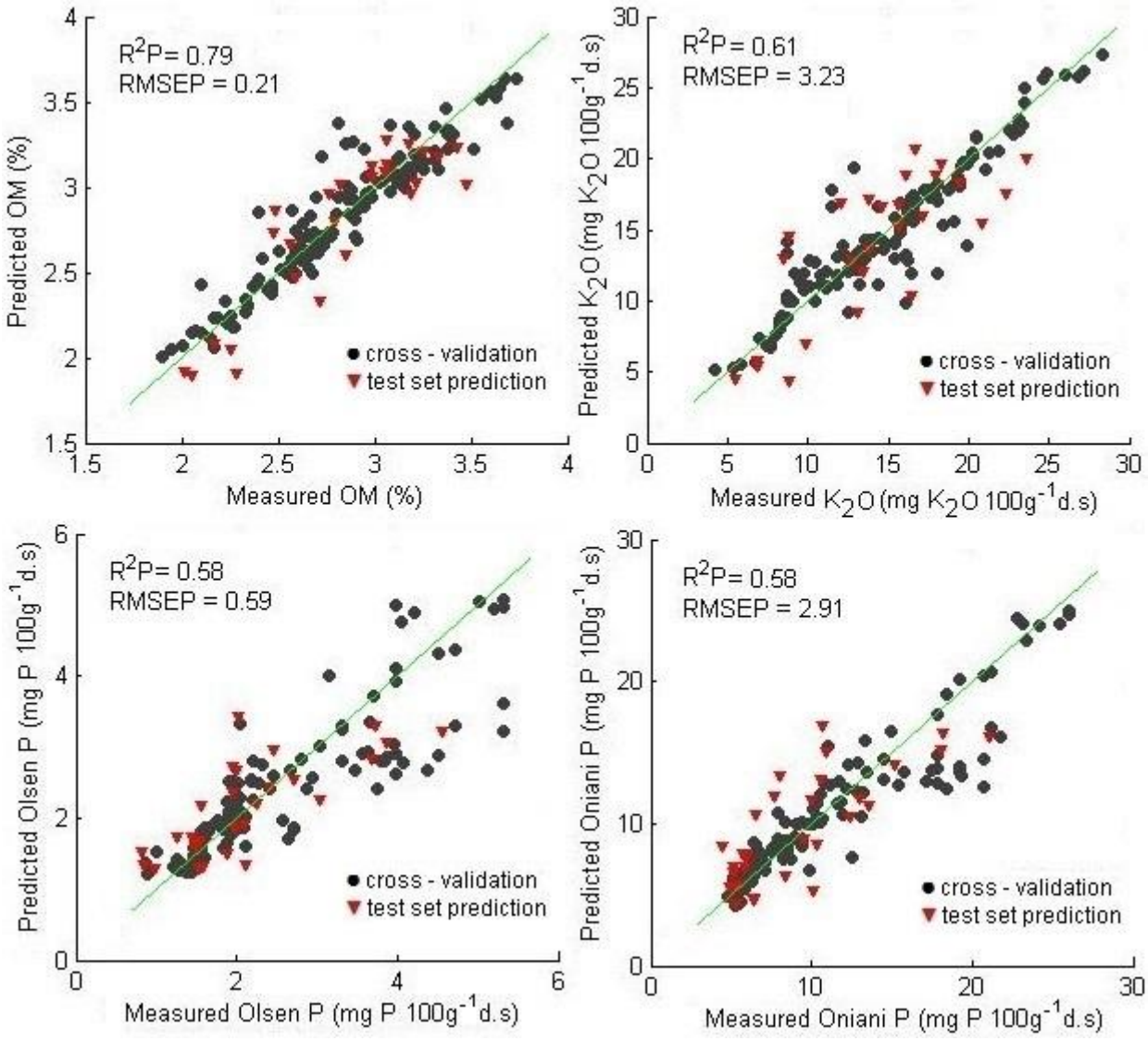


Figure 7-5. Vis/NIR predicted versus chemical conventional analyses of soil fertility parameters within a field on Vertisol soil; OM (SVM), K_2O (SVM), Olsen P (SVM), Oniani P (SVM).

The prediction statistics for the Vertisol soil samples using the three regression models (PLS, LWR and SVM) are listed in Table 7-3. The lower difference among the statistics RMSEC and RMSECV with a difference of 0.01 was observed in OM. All the regression models gave similar results in cross-validation and test set prediction but the RMSEP values in all the cases were best for the non-linear SVM regression model.

Based on the guidelines for the accuracy performance of the regression models used in this research, the predictive accuracy was moderately successful for OM and less reliable for the rest of the soil fertility parameters. However, in Olsen P the model was classified as less reliable (R^2 of 0.58), but the RMSEP (0.59%) values are quite comparable to the RMSECV values (0.62%), for that, this model could be useful.

These results are in line with the reported by Volkan et al. (2010) for air-dried soil samples from an Entisol. They also obtained the best prediction results for organic matter (R^2= 0.76) and the worst for K (R^2= 0.32) by using PLS regression method.

In agreement with Clark et al. (1990) and Clark (1999) soil OM expresses broad absorption peaks in the visible range, which are dominated by chromospheres (spectrally active groups e.g. Fe, OH^- in water and minerals, CO_3^{2-}, Al^{2+}, SO_4^{2-} in minerals) and the darkness of humic acid, and absorption peaks in the NIR range (700–2500 nm) from the overtones and combination bands of O–H, C–H, and N–H. Morón and Cozzolino (2007) used two reference methods for measuring P content (Bray P and Resins P) and achieved similar results in prediction (R^2= 0.58 for Bray-P and R^2= 0.61 for Resins-P).

Table 7-3

Calibration and prediction statistics of the calibration models within a field on Vertisol soil

Soil fertility parameter	Regression model	R^2 C [a]	R^2 CV [b]	R^2 P [c]	RMSEC [d]	RMSECV [e]	RMSEP [f]	RPD [g]	RER [h]	C Bias [i]	CV Bias [j]	P Bias [k]
OM	PLS	0.84	0.83	0.80	0.18	0.17	0.21	1.94	7.03	1.78	-5.92	-0.02
	LWR	0.90	0.87	0.79	0.13	0.15	0.24	1.90	6.89	-0.01	-0.01	-0.09
	SVM	**0.89**	**0.87**	**0.79**	**0.14**	**0.15**	**0.21**	**1.96**	**7.13**	**0.00**	**0.01**	**-0.04**
K_2O	PLS	0.84	0.84	0.60	2.06	2.11	3.03	1.52	5.99	7.10	-0.01	-0.28
	LWR	0.87	0.85	0.61	1.88	2.04	3.37	1.38	5.42	-0.06	-0.09	-0.49
	SVM	**0.87**	**0.86**	**0.61**	**1.87**	**1.98**	**3.23**	**1.43**	**5.62**	**-0.10**	**-0.05**	**-0.31**
Olsen P	PLS	0.73	0.71	0.51	0.59	0.61	0.78	1.19	4.83	-1.33	0.00	-0.12
	LWR	0.81	0.74	0.56	0.50	0.57	0.65	1.42	5.73	0.06	0.06	0.04
	SVM	**0.78**	**0.72**	**0.58**	**0.54**	**0.62**	**0.59**	**1.56**	**6.31**	**-0.11**	**-0.12**	**0.02**
Oniani P	PLS	0.83	0.82	0.56	2.37	2.44	4.23	1.09	4.07	1.78	-0.01	-1.13
	LWR	0.88	0.84	0.57	1.99	2.28	3.50	1.28	4.78	0.32	0.31	-0.44
	SVM	**0.88**	**0.84**	**0.58**	**2.04**	**2.35**	**2.91**	**1.55**	**5.77**	**-0.35**	**-0.34**	**-0.44**

[a] R^2 C- coefficient of determination of calibration; [b] R^2 CV- coefficient of determination of cross validation; [c] R^2 P- coefficient of determination of prediction
[d] RMSEC- root mean square error of calibration; [e] RMSECV- root mean square error of cross validation; [f] RMSEP- root mean square error of prediction;
[g] RPD- ratio of performance deviation; [h] RER- ratio of error range;
[i] C Bias- systematic deviation of calibration; [j] CV Bias- systematic deviation of cross validation; [k] P Bias- systematic deviation of prediction
Data in bold- Best regression model based on the R^2, RMSECV, RMSEP, RPD, RER

7.3.4. Identification of effective wavelength intervals within a field on Vertisol soil

Those wavelengths most significant for predicting soil fertility parameters were identified in Figure 7-6. Within a field, as well as at landscape level, the wavelengths included in the interval number 1 (399 – 489 nm) were selected for OM, Olsen P and Oniani P. Also, there were coincidences as to the three intervals selected for OM [1; 2 (489 – 580 nm); 12 (1402 – 1480 nm)] and in two of the intervals selected for K_2O [2; 3 (580 – 674 nm)], Olsen P [1; 8 (1050 – 1145 nm)] and for Oniani P [1 and 2]. These results showed that similar wavelengths can be identified within a field and at landscape level on Vertisol soil.

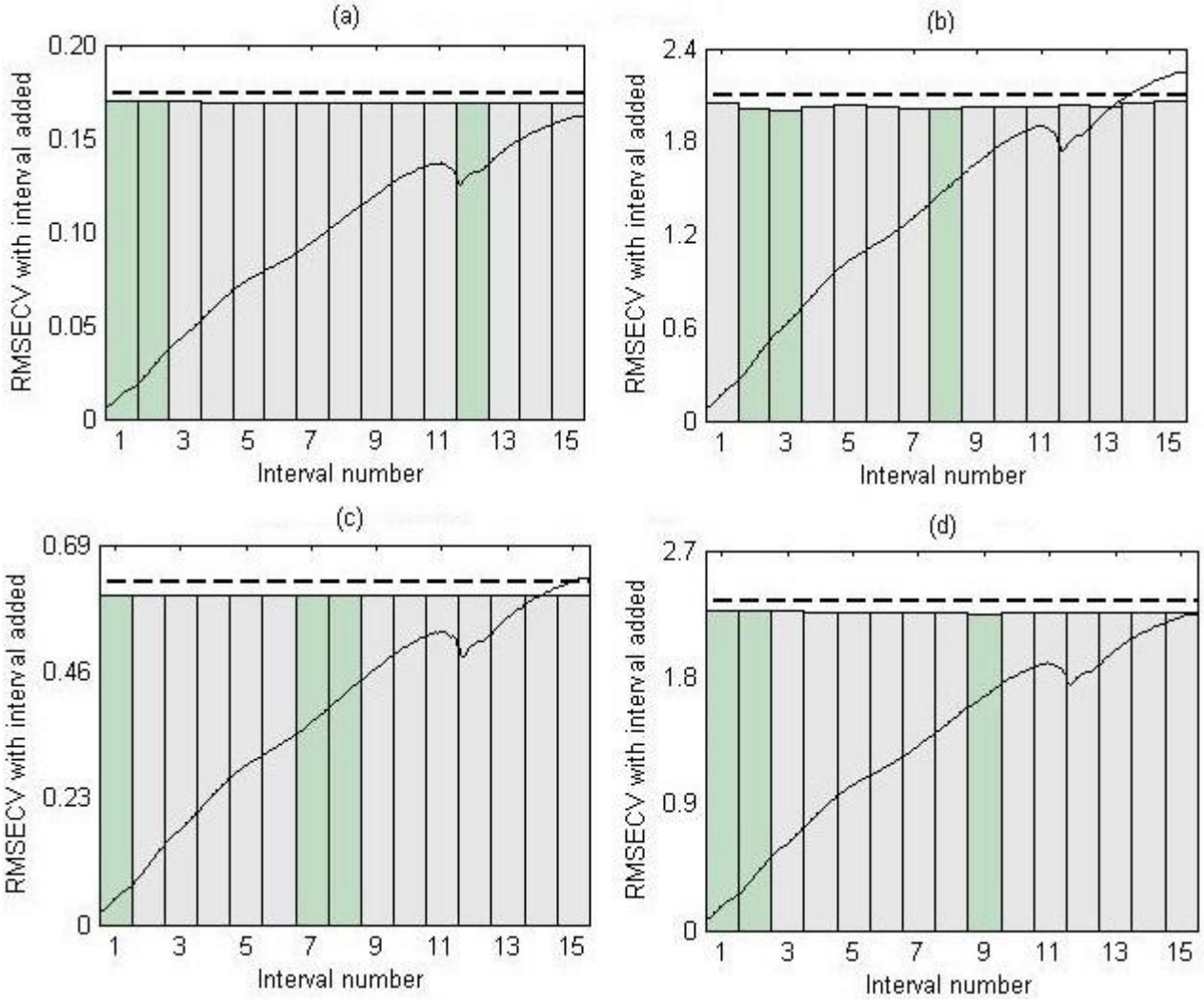


Figure 7-6. Spectral region selected (dark colour) by Forward Interval PLS model (iPLS) within a field on Vertisol soil.

Intervals selected:

(a) OM- [1 (399 – 489 nm), 2 (489 – 580 nm), 12 (1402 – 1480 nm)]
(b) K_2O- [2 (489 – 580 nm), 3 (580 – 674 nm), 8 (1050 – 1145 nm)]
(c) Olsen P- [1 (399 – 489 nm), 7 (952 – 1050 nm), 8 (1050 – 1145 nm)]
(d) Oniani P- [1 (399 – 489 nm), 2 (489 – 580 nm), 9 (1145 – 1235 nm)]

In line with these results Wang et al. (2013) found the best fitness for predicting OM at 399 nm and 449 nm. For Vertisol soil, one of the wavelengths selected by Nowkandeh et al. (2013) for detecting OM (477 nm) is in agreement in both scenarios (landscape level and within a field). This wavelength is included in the interval number 1. The wavelengths selected for K_2O, Olsen P and Oniani P are in line with the reported by other authors. Morón and Cozzolino (2007) performed the PLS loadings for the Vis/NIR calibrations of P in soil samples of agricultural fields from Uruguay, and obtained the highest loadings around the VIS region 500 nm, 550 nm, 600 nm and 660 nm. Thomasson et al. (2001) and Lee et al. (2003) reported wavelengths which are in agreement in both scenarios too. For instance, there was a coincidence with the wavelengths selected for P at 428 nm, 430 nm, 475 nm, 1075 nm, 1100 nm, 1125 nm and 1225 nm. According to the results NIR spectral range can also provide the information needed to quantify these soil fertility parameters.

Table 7-4 compares the results of the predictive regression models within a field based on the selected wavelength intervals. The best regression model (SVM) developed for OM achieves a moderately successful prediction performance (R^2= 0.82). This result is better than the obtained with the full spectrum (moderately useful) with a lower RMSEP value (0.18). Both results are in line with those obtained at landscape level. By reducing the number of wavelengths, the prediction accuracy and the RMSEP values are improved in the three regression models.

By using the 20% of the wavelengths, the soil fertility parameters K_2O, Olsen P and Oniani P increased the R^2 of prediction, reduced the RMSEP and retain the model selected as the best with the full spectrum (SVM). However, the prediction accuracy was the same as observed with the full spectrum ($R^2 < 0.70$ – less reliable).

Table 7-4

Comparison between the predictive regression models within a field on Vertisol soil based on the selected wavelength intervals

Soil fertility parameter	Regression model	R^2 C [a]	R^2 CV [b]	R^2 P [c]	RMSEC [d]	RMSECV [e]	RMSEP [f]	RPD [g]	RER [h]	C Bias [i]	CV Bias [j]	P Bias [k]
	PLS	0.86	0.86	0.79	0.16	0.16	0.21	2.08	7.57	-0.00	0.00	-0.08
OM	LWR	0.92	0.88	0.80	0.12	0.15	0.23	1.88	6.82	-0.00	0.00	-0.08
	SVM	**0.93**	**0.85**	**0.82**	**0.11**	**0.17**	**0.18**	**2.25**	**8.18**	**0.01**	**0.01**	**0.01**
	PLS	0.87	0.86	0.59	1.95	1.98	2.86	1.61	6.32	-0.00	-0.00	0.05
K_2O	LWR	0.88	0.87	0.62	1.85	1.95	2.97	1.56	6.15	-0.04	-0.08	0.44
	SVM	**0.89**	**0.88**	**0.63**	**1.78**	**1.84**	**2.96**	**1.56**	**6.13**	**-0.02**	**-0.01**	**0.28**
	PLS	0.77	0.77	0.52	0.59	0.60	0.75	1.24	5.00	-0.00	-0.00	-0.10
Olsen P	LWR	0.84	0.81	0.61	0.51	0.54	0.65	1.44	5.82	0.05	0.06	0.12
	SVM	**0.84**	**0.79**	**0.59**	**0.50**	**0.57**	**0.59**	**1.57**	**6.34**	**-0.08**	**-0.09**	**0.06**
	PLS	0.85	0.84	0.56	2.46	2.54	3.74	1.22	4.54	0	0.00	-0.79
Oniani P	LWR	0.90	0.89	0.63	2.02	2.19	3.08	1.45	5.41	0.30	0.32	0.27
	SVM	**0.91**	**0.88**	**0.61**	**1.93**	**2.25**	**2.75**	**1.62**	**6.05**	**-0.36**	**-0.33**	**0.19**

[a] R^2 C- coefficient of determination of calibration; [b] R^2 CV- coefficient of determination of cross validation; [c] R^2 P- coefficient of determination of prediction
[d] RMSEC- root mean square error of calibration; [e] RMSECV- root mean square error of cross validation; [f] RMSEP- root mean square error of prediction;
[g] RPD- ratio of performance deviation; [h] RER- ratio of error range;
[i] C Bias- systematic deviation of calibration; [j] CV Bias- systematic deviation of cross validation; [k] P Bias- systematic deviation of prediction
Data in bold- Best regression model based on the R^2, RMSECV, RMSEP, RPD, RER

7.4. Conclusions

Linear and non-linear regression models were implemented to extract characteristic information from spectral data and consequently to establish predictive models. In general, results within a field indicated that all the regression models (PLS, LWR and SVM) provided good correlations between soil spectra and OM. The SVM regression model gave the best results for OM (R^2 of 0.92) and Olsen P (R^2 of 0.83). The best correlation between soil spectra and Oniani P was obtained with LWR (R^2 of 0.72). These better predictions obtained for phosphorus content (Olsen and Oniani) could be related to their higher correlations with OM. The lower prediction R^2 values for K_2O (R^2 of 0.63 with LWR) might be explained the lack of spectral activity of this molecule. Results from the Vis/NIR predictions in Cambisol were much stronger than the results drawn from Vertisol (OM- R^2 of 0.79; K_2O- R^2 of 0.61; - Olsen P-R^2 of 0.58; Oniani P- R^2 of 0.58). The high prediction accuracy for OM, in comparison with the rest of soil fertility parameters, suggests that Vis/NIR spectroscopy could be applicable as a rapid and portable method to determine the OM content in Cambisol and Vertisol soils.

Moreover, regression analyses were completed for each soil fertility parameter by using only three spectral intervals selected by PLS method. Within a field on Cambisol soil there were some wavelengths in common among all soil fertility parameters (399 – 489 nm). The same interval was common for OM, Olsen P and Oniani within a field on Vertisol soil. Data from other intervals would provide similar levels of accuracy. Therefore, verification would be an important additional step toward to improve this methodology. Further study will focus on the improvement of the Vis/NIR spectroscopy methodology considering different soil types.

Chapter 8

Usefulness of Vis/NIR for mapping the soil fertility variation within a field

8.1. Introduction

Wet chemistry extractions for soil analyses are often labor intensive, complicated or environmentally hazardous. The extractions are generally tedious, typically involve expensive equipment, require considerable laboratory space, and normally involve the use of numerous chemical reagents. These factors render contemporary soil analysis methods prohibitively expensive for all but the most serious agronomic pursuits (Hanks, 1999). Therefore, by replacing the wet chemical analyses by Vis/NIR spectroscopic scanning in the laboratory it could be possible to speed up the analyses. Thus, no reagents are required and costs are minimised.

On the other hand, small-scale spatial information about the current distribution of soil fertility parameters is therefore important to make appropriate decisions on the amounts of fertiliser that should be added to the soil. Precision agriculture involves the use of sensor technologies for mapping the spatial variation in soil fertility parameters. The output of these technologies is useful information for a site-specific fertilisation management (Schirrmann et al., 2012; Tekin et al., 2013).

Based on the previous considerations the feasibility of Vis/NIR spectroscopy technique for the prediction of soil fertility parameters and soil testing in laboratory conditions has been evaluated in this study. For that reason the chapter aims to compare the spatial variability and distribution of the soil fertility parameters within a field on Cambisol and Vertisol soil by mapping the results of both wet chemical analyses and the Vis/NIR spectroscopic measurements. Also, the prediction accuracy for K_2O and P based on an estimated relation with OM was evaluated. Finally, the profitability of using conventional and Vis/NIR methods under laboratory for soil testing is also discussed.

8.2. Materials and methods

The practical usefulness of the regression models obtained in laboratory conditions was evaluated by constructing maps for both the measured and the predicted soil fertility values. A kriging interpolation algorithm implemented in the software MATLAB 7.9 (R2009b, The Mathworks) was used for mapping these soil fertility parameters within a field on Cambisol and Vertisol soils. Kriging estimation is a weight based calculation and the predictions are better starting from regular samples than from clustered samples.

For Kriging estimation, the collected data should be spatially correlated. A network of points with a variety of separation distances and in a variety of directions was used. It was not particularly useful to have the data locations equally spaced. Maps were comparing by means of the coefficient of variation (CV) of measured and predicted values. For assessing the prediction accuracy for K_2O and P based on an estimated relation with OM, the scatter plots, which include an equation for the correlation between the variables and R^2, were obtained in Matlab. Different statistics (Minimum, Mean, Maximum values, SD, and CV) for comparing the results were used.

The profitability of Vis/NIR spectroscopy for soil fertility mapping was estimated based on the cost of reagents used as compared with the conventional method.

8.3. Results and discussion

8.3.1. Soil fertility parameter maps and prediction accuracy within a field

In Figures 8-1 and 8-2 the interpolated maps of the wet chemically measured and Vis/NIR predicted values of the soil fertility parameters for the Cambisol field are presented. Moreover, as the prediction performance is less accurate, the CV between the wet chemically measured and Vis/NIR predicted maps increases.

For OM (R^2= 0.92- successful accuracy) the maps are fairly similar, according to the distribution of different colours, which indicates that the distribution of the soil fertility content has been quite well captured by the OM values predicted based on the Vis/NIR spectra. From the CV limits used in this research OM showed a lower variability in both maps, with a similar CV of 12.16% and 12.34% for measured and Vis/NIR predicted values respectively. In the case of K_2O (R^2= 0.63- less reliable accuracy), a significant difference can be observed between the measured (CV of 37.68%) and Vis/NIR predicted (CV of 19.88%) values.

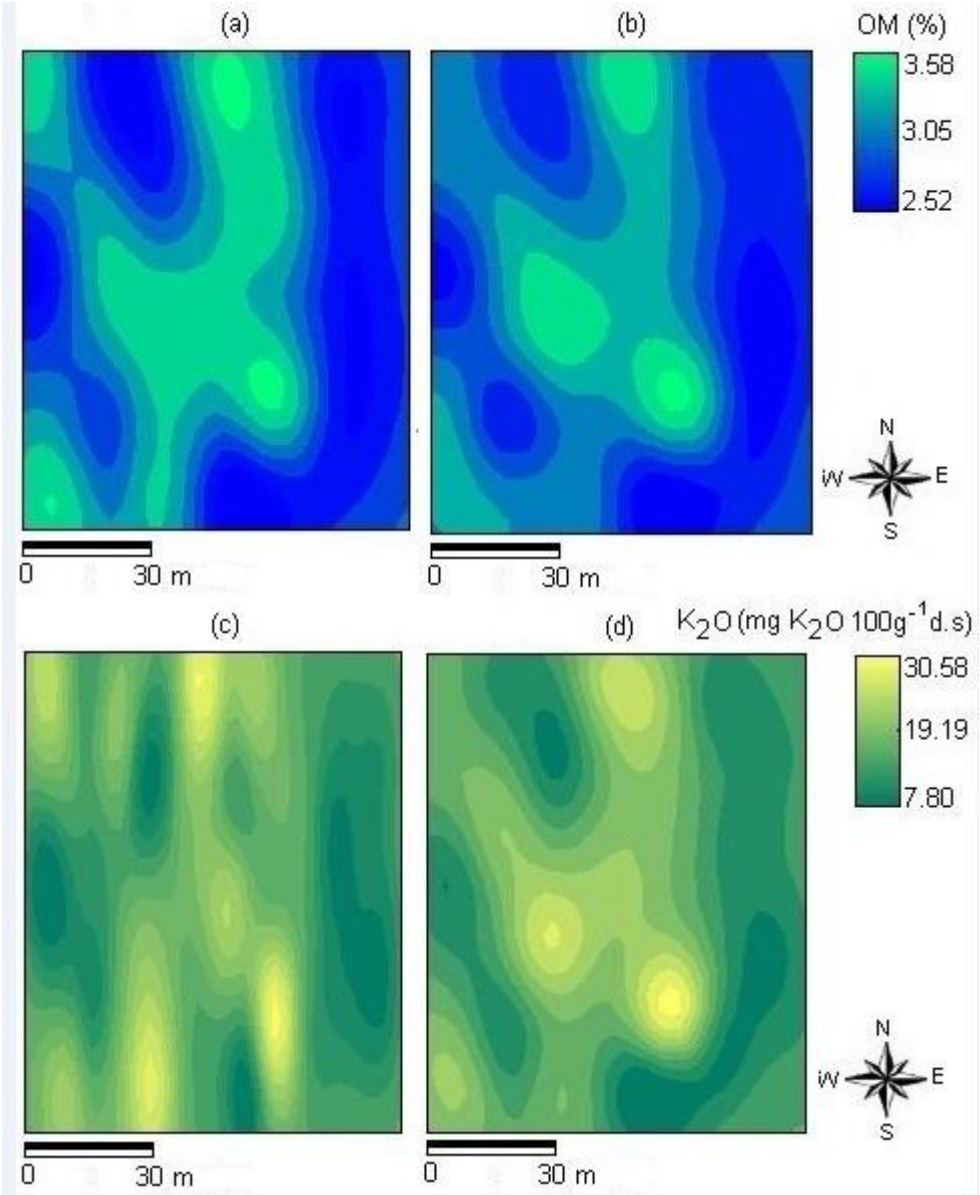


Figure 8-1. Comparison of wet chemically measured (left) and Vis/NIR predicted (right) soil fertility parameters within a field on Cambisol soil.
(a) measured OM, (b) predicted OM, (c) measured K_2O, (d) predicted K_2O).

Coefficients of variation (CV) for wet chemically measured values (OM- 12.16%, K_2O- 37.68%)
Coefficients of variation (CV) for Vis/NIR predicted values (OM- 12.34%, K_2O- 19.88%)

The maps obtained for Olsen P (R^2= 0.83- moderately successful accuracy) are also comparable. In this case, the CV values (15.27% for measured and 12.34% for Vis/NIR predicted values) indicate a lower variability. Finally, the measured and Vis/NIR predicted values for Oniani P (R^2= 0.72- moderately useful) showed a moderate variability, but the differences between them were higher than the observed for OM and Olsen P. This variability exhibited a CV of 31.70% and 26.70% for the measured and Vis/NIR predicted values respectively.

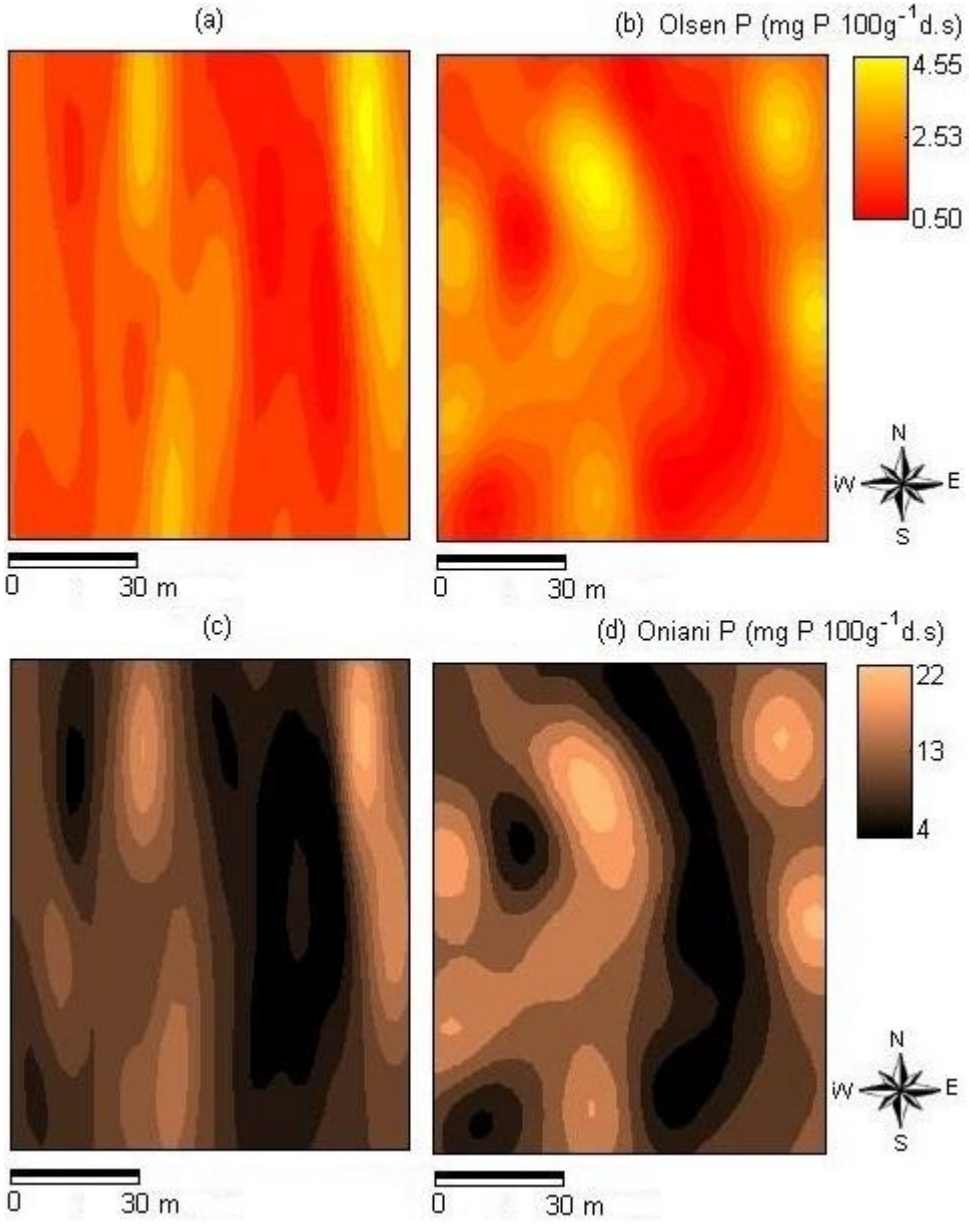


Figure 8-2. Comparison of wet chemically measured (left) and Vis/NIR predicted (right) soil fertility parameters within a field on Cambisol soil.
(a) measured Olsen P, (b) predicted Olsen P, (c) measured Oniani P, (d) predicted Oniani P.
Coefficients of variation (CV) for wet chemically measured values (Olsen P- 15.27%, Oniani P- 31.70%)
Coefficients of variation (CV) for Vis/NIR predicted values (Olsen P- 12.34%, Oniani P- 26.70%)

The maps developed for the measured and predicted values of OM, K_2O, Olsen P, and Oniani P in Vertisol field are illustrated in Figures 8-3 and 8-4. According to the distribution of different colours, these measured and predicted maps showed a good agreement for OM (R^2= 0.79- Moderately useful). This is sustained by the fact that in this soil fertility parameter a lower variability in measured values (14.41%) and in predicted values (16.69%) was observed. Also, in K_2O (R^2= 0.61- Less reliable) the CV values were very similar and corresponded to a moderate variability (32.57% and 34.82% for measured and predicted values respectively).

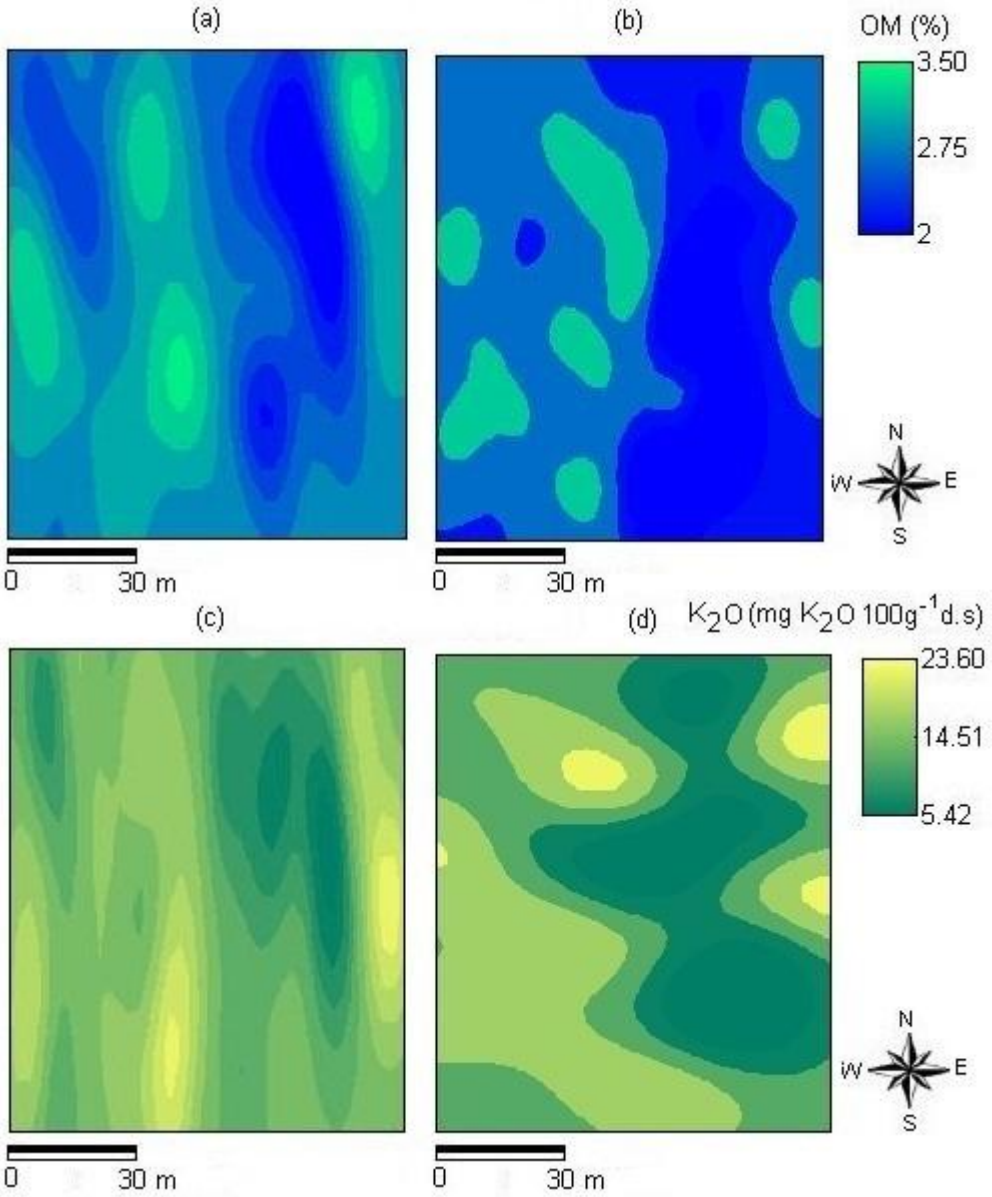


Figure 8-3. Comparison of wet chemically measured (left) and Vis/NIR predicted (right) soil fertility parameters within a field on Vertisol soil.
(a) measured OM, (b) predicted OM, (c) measured K_2O, (d) predicted K_2O).
Coefficients of variation (CV) for wet chemically measured values (OM- 14.41%, K_2O- 32.57%)
Coefficients of variation (CV) for Vis/NIR predicted values (OM- 16.69%, K_2O- 36.65%)

Although the measured and predicted values for Olsen P and Oniani P showed a moderate variability, the differences between them were higher in both parameters than those observed for OM and K_2O. For Olsen P (R^2= 0.58- Less reliable) the CV for the measured values was 43.95%, while for the predicted values it was only 32.40%. For Oniani P (R^2= 0.58- Less reliable) this variability exhibited a CV of 47.02% and 39.91% for the measured and predicted values respectively. This result indicates that both predictive models were less able to detect the variation for phosphorus than for OM and K_2O. These maps also, show differences in various areas.

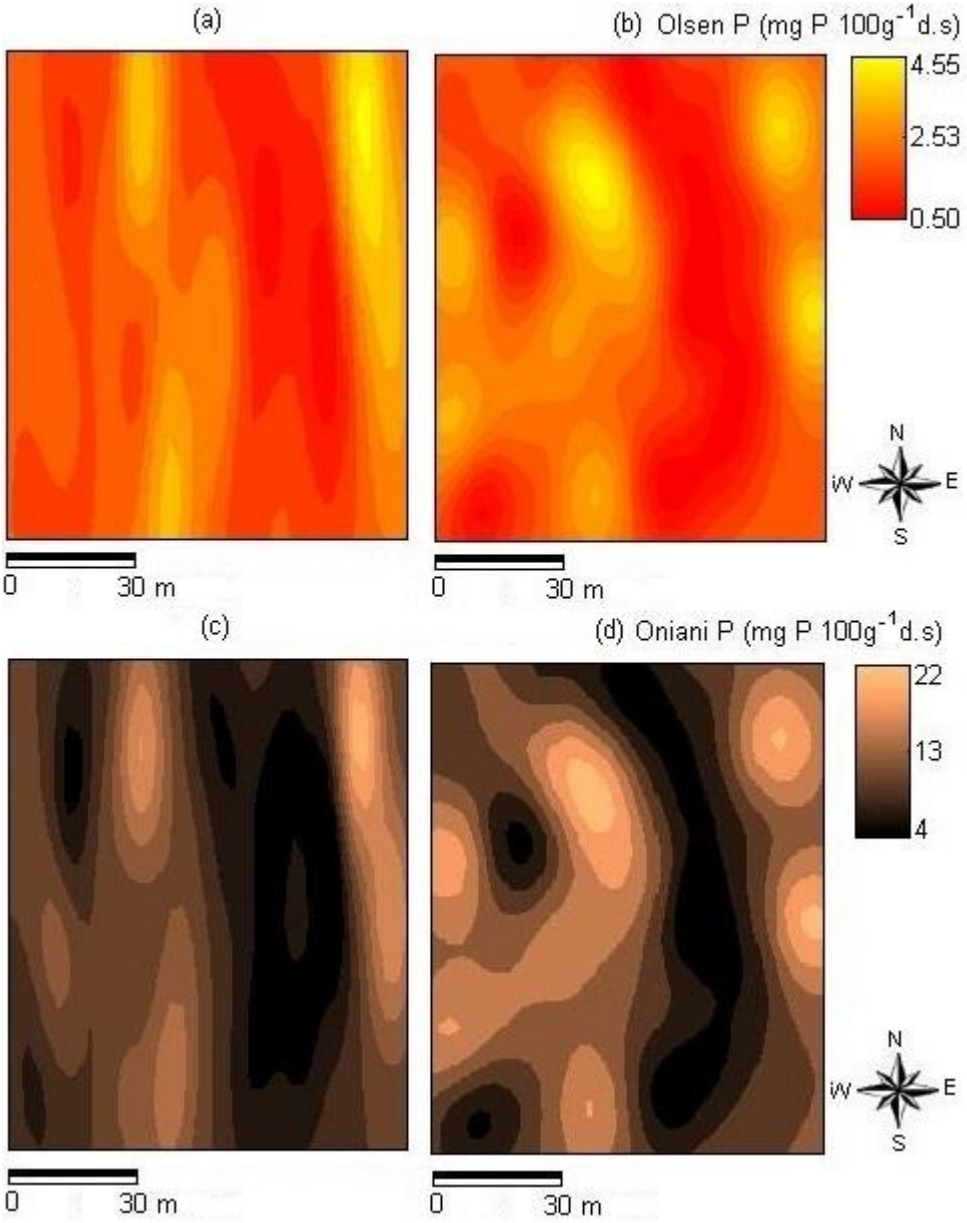


Figure 8-4. Comparison of wet chemically measured (left) and Vis/NIR predicted (right) soil fertility parameters within a field on Vertisol soil.
(a) measured Olsen P, (b) predicted Olsen P, (c) measured Oniani P, (d) predicted Oniani P.
Coefficients of variation (CV) for wet chemically measured values (Olsen P- 43.95%, Oniani P- 47.02%)
Coefficients of variation (CV) for Vis/NIR predicted values (Olsen P- 32.40%, Oniani P- 39.91%)

8.3.2. Prediction accuracy for K_2O and P based on an estimated relation with OM

Figure 8-5 shows the results from the scatter plots corresponding to the prediction accuracy for K_2O and P based on an estimated relation with OM within a field on Cambisol soil.

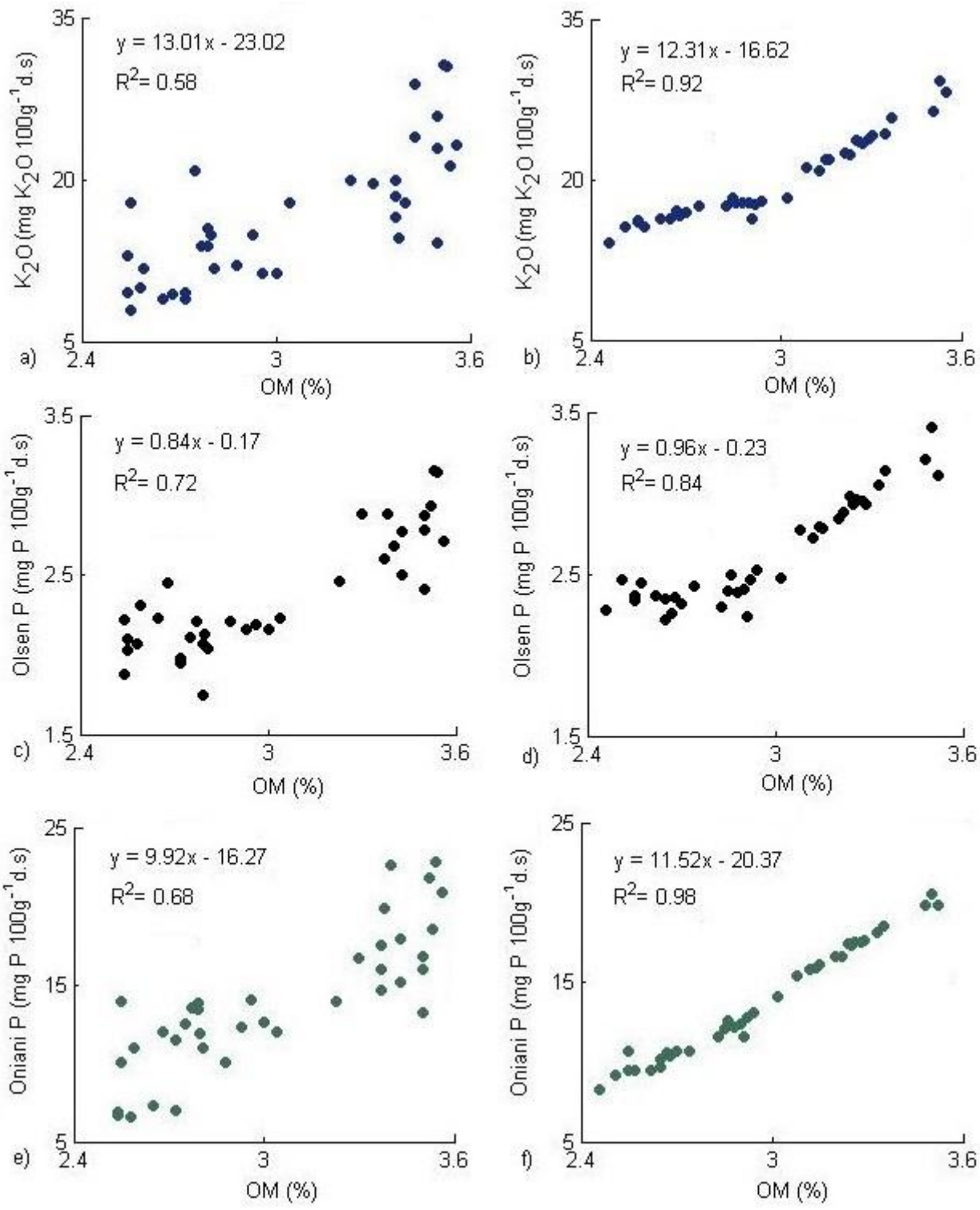


Figure 8-5. Prediction accuracy within a field on Cambisol soil for K_2O and P based on the correlation with OM.
(a) K_2O & OM reference content, (b) K_2O & OM Vis/NIR predicted content
(c) Olsen P & OM reference content, (d) Olsen P & OM Vis/NIR predicted content
(e) Oniani P & OM reference content, (f) Oniani P & OM Vis/NIR predicted content

For all the soil fertility parameters, minor relationships (R^2) were observed between the OM content and the values obtained by means of the reference methods. However, the correlation between the OM content and the other soil fertility parameters as predicted from the Vis/NIR spectra was considerably higher $R^2 \geq 0.84$. This indicates that the K_2O content and the Olsen and Oniani P content in the soil are mostly predicted based on their (weak) correlation with the OM content.

Most of the descriptive statistics, derived from the prediction accuracy for K_2O and P based on an estimated relation with OM in Cambisol field, are illustrated in Table 8-1. These data show that in each soil fertility parameter, the mean values obtained from the predicted values of OM, are consistent with the values obtained from the relation with the reference OM. However, the coefficient of variation (CV) in K_2O showed a higher difference between the values obtained from the measured and Vis/NIR predicted OM. As a consequence, the CV of K_2O obtained from the reference OM indicated a moderate variability (29.05%), while the CV of K_2O obtained from the predicted OM showed lower variability (19.12%). These results are consistent with the lower correlation between the K_2O content and the OM content as measured with the reference methods (R^2= 0.58) than for the P content (R^2= 0.68 – 0.72). This is in line with the findings in Chapter 5 where we found that Vis/NIR spectroscopy is not able to reliably predict the variation in the P content in soil samples which is uncorrelated to the OM content.

Table 8-1

Prediction accuracy for K_2O and P based on an estimated relation with OM within a field on Cambisol soil

Soil fertility parameter	Mean	Min [a]	Max [b]	SD [c] (±)	CV [d] (%)
K_2O (OM_R)	16.56	10.03	23.30	4.81	29.05
K_2O ($OM_{V/N}$)	19.93	13.66	26.73	3.81	19.12
Olsen P (OM_R)	2.39	1.96	2.82	0.31	13.02
Olsen P ($OM_{V/N}$)	2.62	2.13	3.15	0.30	11.34
Oniani P (OM_R)	13.91	8.93	19.05	3.67	26.37
Oniani P ($OM_{V/N}$)	13.83	7.97	20.20	3.57	25.77

[a]Min- minimum observed value; [b]Max- maximum observed value;
[c]SD- standard deviation in the reference data; [d]CV- coefficient of variation
K_2O (mg K_2O 100 g^{-1} d.s); Olsen P (mg P 100 g^{-1} d.s); Oniani P (mg K_2O 100 g^{-1} d.s);
OM_R (%) - Organic matter determined by a reference method (Walkley-Black);
$OM_{V/N}$ (%) - Organic matter predicted by Vis/NIR spectroscopy

The correlations for K_2O and P based on an estimated relation with OM within a field on Vertisol soil are presented in Figure 8-6.

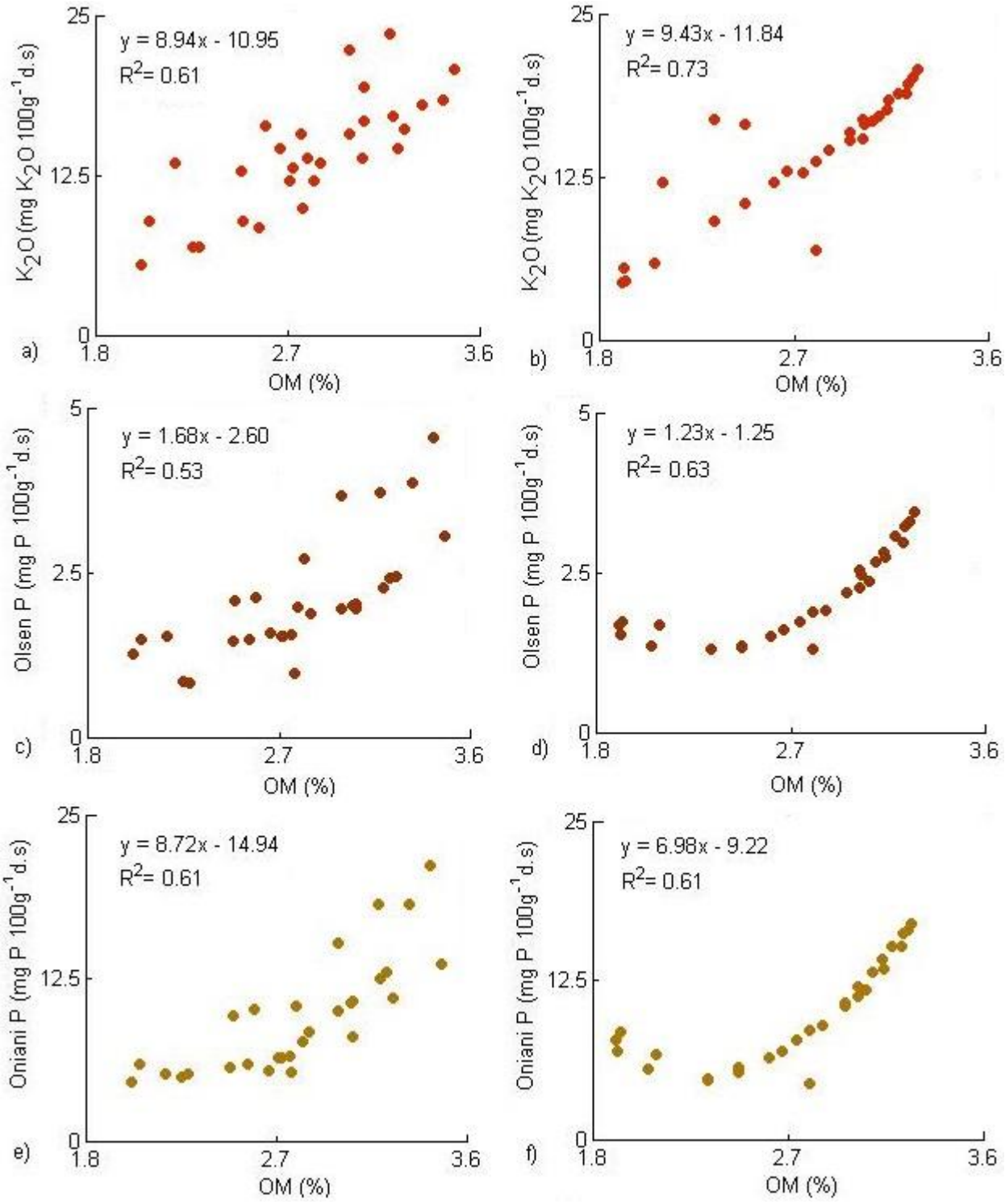


Figure 8-6. Prediction accuracy within a field on Vertisol soil for K_2O and P based on the correlation with OM.
(a) K_2O & OM reference content, (b) K_2O & OM Vis/NIR predicted content
(c) Olsen P & OM reference content, (d) Olsen P & OM Vis/NIR predicted content
(e) Oniani P & OM reference content, (f) Oniani P & OM Vis/NIR predicted content

The scatter plots show lower relationships (R^2) between these paramateres than observed in Cambisol, for both measured and Vis/NIR predicted values. However, the R^2 from the soil fertility parameters shows larger values in the Vis/NIR prediction than those observed in the reference content, except for Oniani P that was equal (R^2 of 0.61) in both stages.

In Table 8-2 the results of the descriptive statistics of K_2O and P based on an estimated relation with OM in Vertisol field are shown. These data demonstrate that the mean values of the predicted values of OM are similar to the values obtained from the relation with the reference OM. The CV values for K_2O and Oniani P indicate the same moderate variability in both fertility parameters, as were obtained from measured and predicted OM. However, in Olsen P the values obtained from the measured OM showed a low variability (CV of 15.57%) and those calculated from the predicted OM showed a moderate variability (CV of 25.91%). In this last case the moderate variability is very close to the limit of 25%.

Table 8-2

Prediction accuracy for K_2O and P based on an estimated relation with OM within a field on Vertisol soil

Soil fertility parameter	Mean	Min [a]	Max [b]	SD [c] (±)	CV [d] (%)
K_2O (OM_R)	13.40	3.13	22.25	5.26	39.26
K_2O ($OM_{V/N}$)	14	6.08	19.04	4.21	30.09
Olsen P (OM_R)	2.18	1.52	2.75	0.34	15.57
Olsen P ($OM_{V/N}$)	2.12	1.09	2.78	0.55	25.91
Oniani P (OM_R)	11.50	3.67	18.25	4.01	34.89
Oniani P ($OM_{V/N}$)	9.91	4.04	13.64	3.12	31.47

[a]Min- minimum observed value; [b]Max- maximum observed value;
[c]SD- standard deviation in the reference data; [d]CV- coefficient of variation
K_2O (mg K_2O 100 g^{-1} d.s); Olsen P (mg P 100 g^{-1} d.s); Oniani P (mg K_2O 100 g^{-1} d.s);
OM_R (%) - Organic matter determined by a reference method (Walkley-Black);
$OM_{V/N}$ (%) - Organic matter predicted by Vis/NIR spectroscopy

8.3.3. Profitability of Vis/NIR spectroscopy for chemical soil testing

The conventional soil analyses in the laboratory employed in this research involves the use of different reagents for determining soil fertility parameters. As illustrated in Table 8-3 these reagents are quite expensive.

Table 8-3

Reagents required on chemical laboratory for soil fertility parameters analysis

Reagent	Formule	Unit size	Unit Price ($)*
Sodium bicarbonate	$NaHCO_3$	1 x 500 g	33.75
Sulfuric acid concentrate	H_2SO_4	12 x 500 ml	45.60
Potassium antimony tartrate hydrate	$C_8H_4K_2O_{12}Sb_2 \cdot xH_2O$	4 x 500 g	186.75
Ammonium molybdate	$(NH_4)_2MoO_4$	1 x 500 g	153.90
Ascorbic acid	$C_6H_8O_6$	1 x 500 g	116.45
Potassium phosphate monobasic	KH_2PO_4	4 x 500 g	72.25
Potassium dichromate	$K_2Cr_2O_7$	4 x 500 g	109

Source: J.T. Baker Chemical Price List
* U.S. dollars

According to the requirements in conventional method for soil chemical analysis, the quantities and cost of reagents are different for the different soil groups (Table 8-4). It should, however, be noted that these costs relate to different number of samples.

Table 8-4

Quantities and costs of reagents used in conventional analysis

Reagent	Quantity [a]		Cost per sample ($)
	Cambisol [b]	Vertisol [c]	
$NaHCO_3$	474.6	277.9	0.14
H_2SO_4	7.7	5.9	3.12
$C_8H_4K_2O_{12}Sb_2 \cdot xH_2O$	0.6	0.4	0.00
$(NH_4)_2MoO_4$	293.8	224.9	0.40
$C_6H_8O_6$	45.9	35.2	0.05
KH_2PO_4	9.9	9.1	0.01
$K_2Cr_2O_7$	221.7	171.6	0.21

[a] Reagent quantities are expressed in g; exception for H_2SO_4 (L);
[b] Reagent quantities are calculated for 226 soil samples;
[c] Reagent quantities are calculated for 173 soil samples.

Table 8-5

Summary of costs for laboratory analysis in conventional method

Items	Cost ($)	%
Reagents [a]	3.93	0.05
Labors [b]	1.33	0.02
Chemical analysis equipment [c]	5 420.00	73
Glassware and supplies [d]	2 000.00	26.93

[a] Reagent costs per sample; [b] Labor costs per sample;
[c] Cost includes the price of the equipment e.g., Laboratory hot plate, Lab-Scale Orbital Shaker, Termo Spectronic, Analytic scale and the cost of the preventative maintainance;
[d] e.g., beakers, erlenmeyer flasks, volumetric flasks, filter paper.

The relatively high cost for laboratory analysis suggests a need for a more powerful tool at a lower cost than using this conventional method. On this sense, the profitability of Vis/NIR spectroscopy according to the cost of using reagents for conventional analysis is demonstrated in Table 8-6 and Table 8-7.

Table 8-6

Profitability of Vis/NIR method as compared to conventional method for chemical soil testing of Cambisol

Method	Reagents	Chemical analysis equipment	Glassware and supplies	Labor costs	Measuring time per sample
			($)		
Conventional	888.18	5 420.00	2 000.00	300.58	1 – 3 h
Vis/NIR	0.00	42 600.00 [a]	158.00 [b]	37.43	< 2 s

[a] Cost includes the price of the mobile diode array spectrophotometer CORONA PLUS REMOTE Vis/NIR and the cost of the preventative maintainance;
[b] Cost includes the price of the petri dish in plastic ($ 0.70 per unit)

Table 8-7

Profitability of Vis/NIR method as compared to conventional method for chemical soil testing of Vertisol

Method	Reagents	Chemical analysis equipment	Glassware and supplies	Labor costs	Measuring time per sample
			($)		
Conventional	679.89	5 420.00	2 000.00	230.09	1 – 3 h
Vis/NIR	0.00	42 600.00 [a]	121.00 [b]	28.65	< 2 s

[a] Cost includes the price of the mobile diode array spectrophotometer CORONA PLUS REMOTE Vis/NIR and the cost of the preventative maintainance;
[b] Cost includes the price of the petri dish in plastic ($ 0.70 per unit)

Therefore, one of the key for using Vis/NIR spectroscopy was to explore the possibility of bringging the cost of the laboratory analysis down to an affordable price in a near future. The lower cost in reagents by investing in new equipment (spectrophotometer) can provide a modern soil analysis method in the Villa Clara province.

On the other hand, equipment maintenance is generally a fundamental factor of laboratory operations. In small laboratories the internal instrument technicians can provide the necessary support due to perform most repairs on devices. However, the cost to repair broken equipment can be considerable, because often additional parts must be acquired from abroad.

8.4. Conclusions

Mapping the measured and predicted soil fertility indicators showed their spatial variability within a field, which is important for site-specific soil fertilising management. These prediction models can be used for variable rate application to maintain an adequate nutrient status for crop production and to potentially reduce the nutrient loss from soils.

At present, fertiliser recommendations for the fields of both soil groups are uniform. However, a soil fertility variation was observed in K_2O (Cambisol) and Olsen P (Vertisol) on the basis of an estimated relation with OM. These results may indicate that there is a need for applying these plant nutrients based on soil nutrient status than a general application for the entire area. However, this is something which must be proven into practice. In general, according to the results obtained from the regression models, this could be more profitability for OM, Olsen P and K_2O in Cambisols and for OM in Vertisols. Consequently, fertiliser application according to the soil status should create the basis for achieving a precisesly fertiliser management in these areas.

Vis/NIR spectroscopy might be effective as a rapid technique for soil fertility assessment without the need for wet-laboratory analysis which involves relative expensive chemical reagents and higher labor costs.

Chapter 9

General conclusions and future work

9.1. General conclusions

Knowledge of spatial variability of soil fertility parameters is important for implementing site specific nutrient management strategies. These strategies can lead to an increase in the effective and efficient use of fertilisers for sugarcane production. A high spatial variability in soil fertility parameters has been reported for Cuban soils. However, previous research has been focusing on the standard soil sampling designs, the conventional chemical laboratory analyses, the determination of an average fertiliser need for the field, and a single rate applied to the entire field. The standard soil analysis procedure is, however, too complex, time-consuming, and expensive to analyse multiple samples within one field as input for site-specific fertilisation. Therefore, the potential of Vis/NIR spectroscopy for simultaneous analysis of multiple soil fertility parameters has been investigated with the aim to develop an efficient tool to be used in the Villa Clara province into the decision-making process for applying precision agriculture in sugarcane production.

Thus, the present dissertation was based on the hypothesis that Vis/NIR spectroscopy could be a useful tool for on-site analyses of soil samples for soil fertility mapping as input to variable rate fertilisation. In order to test this hypothesis the potential of Vis/NIR spectroscopy in laboratory conditions for prediction of soil fertility parameters on dried samples from agricultural fields of Villa Clara province was investigated. Subsequently, to accomplish this main objective three sub-objectives aimed to test the accuracy of Vis/NIR spectroscopy for measuring the total P level in soil when it is not correlated to the organic matter content; to assess the profitability of Vis/NIR spectroscopy for measuring the average fertility values for soil samples at landscape level in the Villa Clara province; and to validate the Vis/NIR spectroscopy technique as a rapid alternative for predicting the variation in soil fertility parameters within a field in the Villa Clara province.

Therefore, the feasibility to develop calibration models based on the correlation between the soil fertility parameters and the acquired Vis/NIR spectra of Cambisol and Vertisol samples from sugarcane fields has been investigated. These models were used for prediction of organic matter, potassium and phosphorus, determined by conventional methods in soil samples from several fields in the same region and for assessing the variation of these parameters within a field.

From the spatial variability and correlation of soil fertility parameters it was concluded that the frequency distribution and range of soil fertility parameters is different for different parameters. Also, the soil fertility, in terms of potassium and phosphorus availability is not low for sugarcane production. However, the large coefficients of variation (CV) observed indicate a large spatial variability in the soil fertility parameters, mainly in phosphorus and potassium content, at the landscape level and even within a field. This means that soil analysis prior to fertilisation is necessary to promote optimal fertilisation.

The variation in the soil fertility parameters, phosphorus and potassium observed within one field corresponded to a CV larger than 30% for both sets from both soil groups demonstrating a moderate variability in these datasets. Only for organic matter, the variation in the present study was considerably lower (CV of 15%). This means that soil analysis prior to fertilisation is necessary, and that this kind of analysis of one mix-sample to obtain the average values for the entire field is not enough for a good nutrient management. Finally, significant correlations were established between all the soil fertility parameters, in general the higher ones were for those related with the organic matter content (0.78; 0.89; 0.88 for K_2O, Olsen P and Oniani P respectively in Cambisol landscape, 0.82; 0.78; 0.83 for K_2O, Olsen P and Oniani P respectively in Vertisol landscape, 0.77; 0.85; 0.83 for K_2O, Olsen P and Oniani P respectively in Cambisol field, and 0.79; 0.74; 0.79 for K_2O, Olsen P and Oniani P respectively in Vertisol field).

The equipment and sample presentation strategies for acquiring these spectra have been described and the soil spectra obtained with this have been discussed. From the acquired reflectance spectra distinct differences were observed between the soil types (Vertisol and Cambisol) and between soil samples from different fields within one soil group. Even within one field clear variation was observed between the acquired reflectance spectra. The reflectance values observed in the Vis region were lower than those in the NIR region.

For soils with added P it was found that all types of soils showed water absorption bands at the same wavelengths around 1400; 1900 and 2250 nm. However, the magnitudes of reflectance of soil spectra for each soil type were different. All types of soils showed water absorption bands at the same wavelengths. The RMSEP and the regression models indicated that the prediction of the added P levels was very inaccurate in the different soil types.

As soil is important for food and feed production and fertilisers are necessary to sustain its fertility, proper type and amount of fertiliser application improves both the quality and yield. With the support of conventional laboratory analyses and soil sampling procedures, it was possible to collect data on the soil fertility status. From Vis/NIR predictions of the average fertility values for soil samples from different fields it was concluded that the soil fertility parameters could be estimated by using non-linear regression models with a successful (OM), moderately successful (Olsen P) and moderately useful (K_2O) accuracy for the Cambisol soil samples. In the case of Vertisol soil samples the non-linear regression models were moderately successful for all considered soil fertility parameters.

The best prediction results were obtained for organic matter in Cambisol at landscape level ($R^2 \geq 0.90$; RMSEP ≤ 0.12) and in Cambisol at field scale ($R^2 \geq 0.89$; RMSEP ≤ 0.14). These results indicate that Vis/NIR spectroscopy could be useful for rapidly determining of different soil fertility parameters. The prediction of phosphorus and potassium was possible through the correlation of these parameters with OM. Consequently, the reliability of the prediction for these parameters depends on the persistence of this correlation with OM. As this correlation is expected to be quite stable over time, adopting the near infrared technique for on-site and even on-line analysis of soil nutrients could provide a very useful tool for farmers in the context of precision agriculture.

The prediction of variation in soil fertility parameters within a field indicated that all the regression models (PLS, LWR and SVM) provided good predictions of the OM content. The SVM regression model gave the best results for both soil groups. The better predictions obtained for phosphorus content in Cambisol could be related to their higher correlations with OM. The low prediction accuracy observed in K_2O and P could be related with the lack of spectral activity of these parameters.

The Vis/NIR prediction accuracy of the soil fertility parameters observed in Vertisol was significantly lower than that obtained in Cambisol, with the exception of K_2O (R^2= 0.63 and R^2= 0.61 for Cambisol and Vertisol respectively). This result might be related to the differences in the correlations between these parameters and the OM content and in the residual soil moisture content. Even after air drying the Vertisol soils contain a considerably amount of water due to the high water-holding capacity of the clay. This remaining soil moisture content has an impact on the reflectance of the air dried soil samples across the electromagnetic spectrum.

The high prediction accuracy for OM suggests that Vis/NIR spectroscopy could be applicable as a measurement method for OM. Mapping the measured and predicted soil fertility indicators showed their spatial variability within a field, which is important for site-specific soil fertilisation management. However, there is a remarkable need for the effective collection of data for precision agriculture, since the soil sampling and laboratory techniques are costly and time consuming. Therefore, utilization of Vis/NIR spectroscopic techniques can improve the efficiency of precision agriculture techniques by giving the chance of making rapid soil analyses within a few seconds.

Consequently, with the use of instrumental techniques, site specific crop management becomes a promising strategy that may be able to increase crop production as well as to reduce input costs and providing environmental sustainability. Site Specific Crop Management is a management technique that seeks to address the variability within a field, and optimize the application of inputs. In this concept, soil fertility can be considered as one of the important soil variables to be sensed and may show variance within a field due to the spatial differences in soil, soil type, previous management practices and agronomic changes.

9.2. Future work

9.2.1. On-line sensing of soil fertility parameters, based on optical reflectance techniques

In contrast with other countries of the region, the use of fertilisers in Cuba, from the beginning of the eighties in the last century, took into account the soil characteristics, particularly the soil nutrient content. However, the fertiliser manufacturing industry of Cuba is no longer operational and all fertilisers are imported. At present Cuba is one of the Latin America countries, which demands a lot of quantities of fertilisers per area. Then, the efficient fertiliser application is important for increasing crop yields and

quality. Therefore, the research related to this topic is essential to bring recommendations on the rational use of fertilisers. In this sense, future works must be focused on the feasibility and possibility of using transformed in-site spectra by calibration technique to predict soil fertility parameters. The Vis/NIR spectroscopy as an efficient tool to predict within-field soil fertility parameters, is significantly valuable when establishing agricultural field trials and in precision farming. Therefore, the on-line measurement of soil fertility parameters is important for site-specific application of fertilisers into agricultural soils. Consequently, a soil sensor for on-line sensing of soil fertility parameters must be developed. This sensor will be attached to a subsoiler, to acquire soil spectra from the bottom of the furrow opened by this implement. The mobile diode array spectrophotometer (CORONA PLUS REMOTE Vis/NIR SB, Zeiss, Jena, Germany), with a measurement range of 399–1697 nm used in this research, will be used to measure soil spectra in reflectance mode. Analysis and interpretation of spatial variability of soils is a key- stone in site-specific farming. Precision agriculture is an information-based management intensive approach to farming. Instead of managing a field as a whole, the philosophy of precision agriculture is to manage individual areas within a field. Accounting for soil variability is a critical need for precision agriculture. Traditionally, quantification and mapping of soil properties have been done through relatively coarse grid soil sampling and statistical interpolation.

9.2.2. Variable rate (VR) fertiliser application

Fertiliser use is not an end in itself. Increased agricultural production and food availability can, in turn, be seen as an objective for the agricultural sector in the context of contributing to expansive the macroeconomic objectives of society. Variable rate (VR) fertiliser application will enable producers to vary the rate of crop fertiliser. This technology combines a VR control system with application equipment to apply inputs at a precise time and location to achieve site-specific application rates of fertiliser. For that reason it will be necessary to develop new calibration models and upgrade previously developed in this research by using Vis/NIR spectroscopy. These regression models might be used for analysing the observed spatial variability in soil fertility parameters. Consequently, it would be possible to propose a VR fertiliser strategy. According to the regression models obtained in this study, this could be more applicable to OM, Olsen P and K_2O in Cambisols and to OM in Vertisols. This strategy could helps to maintain

proper levels of these soil fertility parameters for sugarcane production and also to reduce the loss of fertilisers.

Also, these models will enable measurement of soil fertility parameters from other important soil types for agricultural production in Villa Clara province. Then, these models will aim to improve fertiliser recommendations on the basis of more soil fertility parameters measurements than conventional chemical analyses. Such measurements will then be used to apply fertiliser or management practices on a VR basis leading to cost savings and reduced environmental impacts such as water contamination through nutrient loss and leaching. This site-specific approach will allow farmers to apply products only where they are needed in a field. Varying the application of fertilisers will reduce the input and labor costs, maximize productivity, and reduce the impact of over-application on the environment.

References

Adamchuk V.I., Hummel J. W., Morgan M.T., Upadhyaya S.K. (2004). On-the-go soil sensors for precision agriculture. *Computers and Electronics in Agriculture*, **44**, 71 – 91.

Adamchuk V.I., Viscarra Rossel R.A. (2010). Development of On-the-Go proximal soil sensor systems. *In: Viscarra Rossel R.A., McBratney A.B., Minasny B. (Eds). Proximal Soil Sensing, Progress in Soil Science*, Springer Netherlands, Netherlands, 15-28.

Adamchuk V. (2011). On-the-go soil sensors — are we there yet? *In: Proceeding of The Second Global Workshop on Proximal Soil Sensing*, Montreal, 160-163.

Adekayode F.O., Aderibigbe A. O., Balogun A. M., Adedayo J. O., Oladimeji J. A. (2009). Use of geospatial technology in a precision soil fertility investigation of a farmland for arable crop production in a tropical environment. *Food, Agriculture & Environment*, **2** (7), 852-855.

Adhikari P., Shukla M. K., Mexal J.G. (2011). Spatial variability of electrical conductivity of desert soil irrigated with treated wastewater: implications for irrigation management. *Applied and Environmental Soil Science*, **2011**, 1-11.

Adrian A.M., Norwood S.H., Mask P.L. (2005). Producer's perceptions and attitudes toward precision agriculture technologies. *Computers and Electronics in Agriculture*, **3** (48), 256 - 271.

Aïchi H., Fouad Y., Walter C. (2009). Regional predictions of soil organic carbon content from spectral reflectance measurements. *Biosystems Engineering*, **3** (104), 442-446.

Akinrinde E.A. (2006). Soils: Nature, fertility conservation and management, 122. AMS Publishing, Inc., Vienna, Austria.

Alonso-Pippo W., Luengo C. A., Koehlinger J., Garzone P., Cornacchia G. (2008). Sugarcane energy use: The Cuban case. *Energy Policy,* **36**, 2163-2181.

Aweto A.O. (1982). Variability of upper slope soils developed under sandstones in South-western Nigeria. *Georg. J.*, **25**, 27-37.

Bah A., Balasundram S.K., Husni M.H.A. (2012). Sensor technologies for precision soil nutrient management and monitoring. *American Journal of Agricultural and Biological Sciences*, **7** (1), 43-49.

Bai Y., Wang Y. (2011). Spatial variability of soil chemical properties in a Jujube Slope on the Loess Plateau of China. *Soil Science*, **176**, 550-558.

Barnes R. J., Dhanoa M. S., Lister S.J. (1989). Standard Normal Variate Transformation and Detrending of Near-Infrared Diffuse Reflectance Spectra. *Applied Spectroscopy*, **43** (5), 772-777.

Barzegar A.R., Asoodar M.A., Ansari M. (2000). Effectiveness of sugarcane residue incorporation at different water contents and the proctor compaction loads in reducing soil compactibility. *Soil and Tillage Research*, **57**, 167–172.

Bell J.C., Cunningham R.L, Havens M.W. (1992). Calibration and validation of a soil–landscape model for predicting soil drainage class. S*oil Science Society of America Journal*, **56**, 1860-1866.

Ben-Dor E., Inbar Y., Chen Y. (1997). The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400–2500 nm) during a controlled decomposition process*. Remote Sensing of Environment,* **61**, 1-15*.*

Ben-Dor E., Irons J.R., Epema G.F. (1999). Soil Reflectance. *In: Rencez A.N. (Ed). Remote sensing for the earth science: Manual of remote sensing*, 3rd ed. John Wiley and Sons, New York, 111-188.

Ben-Dor E. (2002). Quantitative remote sensing of soil properties. *Advances in Agronomy*, **75**, 173-243.

Ben-Dor E., Patkin K., Banin A., Karnieli A. (2002). Mapping of several soil properties using DAIS-7915 hyperspectral scanner data: a case study over clayey soils in Israel. *International Journal of Remote Sensing*, **23** (6), 1043-1062.

Bingham F.T. (1962). Chemical soil tests for available phosphorus. *Soil Science*, **94**, 87-95.

Birth G.S., Hecht H.G. (1987). The physics of near infrared reflectance, *In*: *Williams P., Norris K. (Eds). Near infrared technology in the agricultural and food industries.* American Association of Cereal Chemists, Inc., 1-6.

Bocheng X., Xuzhang X., Jihua W., Guodong W. (2004). Spectral characteristics of brown-humid soil and estimation of soil organic matter content by soil reflectivity. *Chinese Journal of Soil Science*, **35** (4), 391–395.

Bogrekci I., Lee W.S. (2005a). Spectral soil signatures and sensing phosphorus. *Biosystems Engineering*, **92** (4), 527-533**.**

Bogrekci I., Lee W.S. (2005b). Improving phosphorus sensing by eliminating soil particle size effect in spectral measurement. *Transactions of the American Society of Agricultural Engineers*, **48**, 1971-1978.

Bogrekci I., Lee W.S. (2005c). Spectral phosphorus mapping using diffuse reflectance of soils and grass. *Biosystems Engineering,* **91** (3), 305-312.

Bogrekci I., Lee W.S. (2007). Comparison of ultraviolet, visible, and near infrared sensing for soil phosphorus. *Biosystems Engineering*, **2** (96), 293-299**.**

Bokobza L. (1998). Near infrared spectroscopy. Jo*urnal of Near Infrared Spectroscopy*, **6**, 3-17.

Borggaard, C. (2001). Neural networks in near-infrared spectroscopy. *In: Williams P., Norris K. (Eds). Near-Infrared Technology in the Agricultural and Food Industries.* Minnesota: American Association of Cereal Chemists, Inc., 101-107.

Bot A., Benites J. (2005). The importance of soil organic matter (Key to drought-resistant soil and sustained food production). Food and Agriculture Organization of the United Nations, Rome, 80.

Braud I., Haverkamp R., Arrue J.L., Lopez MV. (2003). Spatial variability of soil surface properties and consequences for the annual and monthly water balance of a semiarid environment (EFEDA Experiment). *Journal of Hydrometeorology*, **4**, 121-137.

Bricklemyer R.S., Brown D.J. (2010). On-the-go VisNIR: Potential and limitations for mapping soil clay and organic carbon. *Computers and Electronics in Agriculture*, **1** (70), 209-216.

Brouder S., Hofmann B., Reetz H.F. (2001) Evaluating spatial variability of soil parameters for input management. *Better Crops,* **85**, 8–11.

Camilli A., Cugnasca C.E., Saraiva A.M., Hirakawa A.R., Correa P.L.P. (2007). From wireless sensors to field mapping: anatomy of an application for precision agriculture. *Computers and Electronics in Agriculture*, **58**, 25-36.

Caravaca F., Lax A., Albaladejo J. (1999). Organic matter, nutrient contents and cation exchange capacity in fine fractions from semiarid calcareous soils. *Geoderma*, **93**, 161-176.

Cécillon L., Barthès B.G., Gomez C., Ertlen D., Genot V., Hedde M., Stevens A., Brun J.J. (2009). Assessment and monitoring of soil quality using near-infrared reflectance spectroscopy (NIRS). *European Journal of Soil Science*, **60** (5), 770-784.

Chang C.W., Laird D.W., Mausbach M.J., Hurburgh C.R. (2001). Near-infrared reflectance spectroscopy—principal components regression analyses of soil properties. *Soil Science Society of America Journal*, **65**, 480-490.

Chizhikova N.P. (2005) Vertisols of Cuba: mineralogical composition and response to agrogenic impacts (plowing, irrigation, and fertilisation). *Eurasian Soil Science*, **10** (45), 1048.

Choudhury S., Deepankar Ch., Choudhury S. (2009). An Introduction to geographic information technology. I.K. International Publishing House Pvt. Ltd., New Delhi, India, 300.

Christy C.D. (2008). Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy. *Computers and Electronics in Agriculture*, **1** (61), 10-19.

Clark R.N., Gallagher A.J., Swayze G.A. (1990). Material Absorption Band Depth Mapping of Imaging Spectrometer Data Using a Complete Band Shape Least-Squares Fit with Library Reference Spectra. *In: Proceedings of the Second Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Workshop*. JPL Publication 90-54, 176-186.

Clark J.S., Steele I.A., Fender R.P., Coe M.J. (1999). A&A, 348, 888, Near-IR spectroscopy of candidate B[e]/X-ray binaries.

Corwin D.L. (2005). Delineating site-specific crop management units: Precision agriculture application. *In: GIS. ESRI International User Conference Proceedings*.

Corwin D.L., Kaffkab S.R., Hopmans J.W., Morid Y., van Groenigen J.W. , van Kesselb C., Lesch S.M., Ostere J.D. (2003). Assessment and field-scale mapping of soil quality properties of a saline-sodic soil. *Geoderma*, **1952**, 1-29.

Corwin D.L., Lesch S.M. (2008). Application of Geo-referenced Geophysical Measurements to Precision Agriculture. *FastTIMES*, **13** (2), 29-37.

Cogdill R.P. (2008). Case for process analytical technology: regulatory and industrial perspectives. *In: Gad S.C. (Ed). Pharmaceutical Manufacturing Handbook*. Hoboken, N.J., Wiley-Interscience, 391-393.

Cox M.S., Gerard P.D., Wardlaw M.C., Abshire M.J. (2003). Variability of selected soil properties and their relationships with soybean yield. *Soil Science Society of America Journal*, **4** (67), 1296-1302.

Cozzolino D., Murray I. (2004). Identification of animal meat muscles by visible and near infrared 281 reflectance spectroscopy. *Lebensmittel-Wissenschaft & Technologie*, **37**, 447- 452.

Cropsci (2012). Precision agriculture. Department of crop science. College of agriculture and life sciences. Available in: http://www.cropsci.ncsu.edu/PA.asp

Dalal R.C., Henry R.J. (1986). Simultaneous determination of moisture, organic carbon and total nitrogen by near infrared reflectance spectrophotometry. *Soil Science Society of America Journal,* **50**, 120-123.

Daniels M., Daniel T., Van Devender K. (2004). Soil phosphorus levels: Concerns and recommendations. Univ. of Arkansas Division of Agriculture, Cooperative Extension Service, Little Rock, AK. Available in: http://www.sera17.ext.vt.edu/Documents/Soil_P_Levels_Concerns_and_Recommendations.pdf

Dardenne P., Sinnaeve G., Baeten V. (2000). Multivariate calibration and chemometrics for near infrared spectroscopy: Which method? *Journal of Near Infrared Spectroscopy*, **8**, 229-237.

Day P.R. (1965). Particle fractioning and particle size analysis. *In: Black CA (Ed). Methods of soil analysis*, part 1. ASA, Madison, WI, 545-562.

De León M.E., Campos J., Betancourt G., Machado I., Soler L. (2004). Generalización de Lotes Controles del Servicio de Recomendaciones de Fertilizantes y Enmiendas, vía para incrementar el rendimiento agrícola de la caña de azúcar. XV FORUM de Ciencia y Técnica, Ciudad de la Habana, July 8, 28.

DeLiberty T. (1990). Geog 474. Energy Interactions with the Atmosphere and at the Surface. Available in: http://www.udel.edu/Geography/DeLiberty/Geog474/geog474_energy_interact.html

Dessipri E., Minopoulou E., Chryssikos G.D., Gionis V., Paipetis A., Panayiotou C. (2003). Use of FT-NIR spectroscopy for on-line monitoring of formaldehyde-based resin synthesis. *European Polymer Journal*, **39**, 1533-1540.

Donigian A.S. (2002). Watershed model calibration and validation: The HSPF experience. Water Environment Federation National TMDL Science and Policy. Specialty Conference, November.

Ehsani R., Schumann A., Salyani M. (2009). Variable Rate Technology for Florida Citrus. EDIS, 5. Available in: http://edis.ifas.ufl.edu/AE444

Eigenvector (2012). Advanced Preprocessing: Variable Centering. Available in: http://wiki.eigenvector.com/index.php?title=Advanced_Preprocessing:_Variable_Centering

EPA (1997). Environmental Fact Sheet: Waste-Derived Fertilisers. USA Environmental Protection Agency. Available in: http://www.epa.gov/osw/hazard/recycling/fertiliz/fertiliz.txt

EPA (2010). Managing agricultural fertiliser application to prevent contamination of drinking water. Water Protection Practices Bulletin. USA Environmental Protection Agency (EPA). Available in: http://www.epa.gov/safewater/sourcewater/pubs/fs_swpp_fertiliser.pdf

Esbensen K.H., Guyot D., Westad F., Houmøller L.P. (2002). Multivariate Data Analyses. *In: Practice, 5th ed.*, Camo, Oslo.

Erickson B. (2004). Field Experience Validates On-The-Go Soil pH Sensor. Purdue Top Farmer Crop Workshop Newsletter, December, 4.

FAO (Ed) (1990). FAO-Unesco soil map of the world. Revised legend. Soils Bulletin 60. Food and Agriculture Organization of the United Nations, Rome.

FAO/ISRIC/UNEP/CIP (1998). Soil and terrain database for Latin America and the Carribean. *FAO Land and Water Digital Media Series, 5.* FAO, Rome.

FAO (2001). Lecture notes on the major soils of the world. World Soil Resources Reports, 94. Food and Agriculture Organization of the United Nations, Rome, 334.

FAO (2003). Fertiliser use by crop in Cuba. Land and Plant Nutrition Management Service Land and Water Development Division. Food and Agriculture Organization of the United Nations, Rome, 28.

Fassbender J.W. (1975). Química de suelos, con énfasis en suelos de América Latina. IICA, Turialba, Costa Rica, 38.

Fischer G., Teixeira E., Hiznyik E.T., Van Velthuizen H. (2008). Land use dynamics and sugarcane production. *In: Zuurbier P., van de Vooren J. (Eds). Sugarcane ethanol - Contributions to climate change mitigation and the environment*. Wageningen Academic Publishers: Wageningen, The Netherlands, 29-62.

FIU (2012). General Information on Cuba. Florida International University. Available in: http://www2.fiu.edu/~fcf/gen.info.cuba.html

Florinsky I.V., Eilers R.G., Manning G.R., Fuller L.G. (2002). Prediction of soil properties by digital terrain modelling. *Environmental Modelling & Software*, **17**, 295-311.

Forment S., Hanselaer P. (2010). Stray light performance of a combined monochromator–spectrograph UV irradiance measuring instrument. *Measurement Science and Technology,* **8** (21).

Gersper P.L., Rodrígues-Barbosa C.S., Orlando L.F. (1993). Soil conservation in Cuba: A key to the new model for agriculture. *Agriculture and Human Values,* **10** (3), 16–23.

Goddu R.F., Delker D.A. (1960). Spectra-structure correlations for the near-infrared region. *Analytical Chemistry*, **32**, 140-141.

Guillén A., Moral F.G., Herrera L.J., Rubio G., Rojas I., Valenzuela O. Pomares H. (2010). Using near-infrared spectroscopy in the classification of white and iberian pork with neural networks. *Neural Computing and Applications,* **3** (19), 465-470.

Guimarães E., Stein A., Klamt E. (1997). Large area spatial variability of soil chemical properties in central Brazil. *Agriculture, Ecosystems and Environment*, **66** (2), 139-152.

Hami O. (1974). Allgemeine Charakterisierung, Phosphat und Kaliumhaushalt verschiedener

Böden Marokkos. Karl Marx Universität. Diss.

Haneklaus S., Schung E. (1998). Impacts of precision agriculture technologies on fertilisation. *In: Proceedings of the 11th International Symposium of CIEC-Codes of Good Fertiliser Practice and Balanced Fertilisation*, 95-107.

Haneklaus S., Schnug E. (2006) Site-specific nutrient management: objectives, current status and future research needs. *In: Srinivasan A. (Ed). Handbook of precision agriculture: principles and applications*. New York: Food Products Press, 9-34.

Hanks D.A. (1999). Soil nutrient extraction method using pressurized hot water. USPTO Assignment. Data provided by IFI CLAIMS Patent Services.

Harris D.C. (2002). Quantitative chemical analysis. *6th ed. Freeman W.H. and Company*, New York.

He Y., Song H. Y., Garcia A., Hernandez A. (2005). A new approach to predict N, P, K and OM content in a loamy mixed soil by using near infrared reflectance spectroscopy. *Lecture Notes in Computer Science*, **3644**, 859-867.

He Y., Huang M., Garcia A, Hernández A., Song H. (2007). Prediction of soil macronutrients content using near-infrared spectroscopy. *Computers and Electronics in Agriculture*, **2** (58)**,** 144-153.

Heise H.M., Winzen R. (2002). Chemometrics in Near – Infrared Spectrsoscopy. *In: Siesler H.W., Ozaki Y., Kawata S., Heise H.M. (Ed). Near-infrared spectroscopy. Principles, instruments, applications*. Wiley-VCH Verlag, Weinheim, 125-162.

Herrera J.A., Rodriguez R., Herrera J.L., Fardeau J.C. (2002). Evaluation of methods for quantifying bioavailable phosphorus in a Ferralsol from Cuba. *In: Assessment of soil phosphorus status and management of phosphatic fertilisers to optimize crop production*, 117-125. Vienna, IAEA TECDOC 1272 IAEA, 473.

Hernández J.A., Ascanio M.O., Morales M., Cabrera A. (2005). Correlación de la nueva Versión de Clasificación Genética de los suelos de Cuba con las clasificaciones internacionales: Una herramienta útil para la investigación, docencia y producción agropecuaria. INCA. MES. CEVIM. Universidad Veracruzana, Mexico. Instituto de Suelos. MINAGRI. La Habana, Cuba, 62.

Hill M.C. (1998) Methods and guidelines for effective model calibration. U.S Geological Survey Water-Resources Investigations Report 98-4005. US Geologicl Survey, Denver, Colorado.

Hodges S.C. (2010). Soil Fertility Basics, Soil Science Extension, North Carolina State University.

Hong S.Y., Sudduth K.A., Kitchen N.R., Drummond S.T., Palm H.L., Wiebold W.J. (2002). Estimating within field variations in soil properties from airborne hyperspectral images. *In: ISPRS Commission I/ FIOEOS 2002 Conference Proceedings*.

Horneck D.A. (2011). World fertiliser demand and supply. *In: Proceedings, 2011 Western Alfalfa & Forage Symposium*, Las Vegas, NV, 11-13 December. UC Cooperative Extension, Plant Sciences Department, University of California, Davis, CA 95616. Available in: http://alfalfa.ucdavis.edu/+symposium/2011/files/talks/11WAS-4_Horneck_FertiliserDemand.pdf

Hummel J.W., Birrell S.J. (1998). Real-time close-range Sensors for soil attributes. *In: Illinois Fertiliser Conference Proceedings*, January 26-28.

Igne B., Reeves J.B., McCarty G.W., Hively W.D., Lund E., Hurburgh C.R. (2010). Evaluation of PLS, LS-SVM, and LWR for quantitative spectroscopic analysis of soils. *Geoderma*, **18** (3), 167-176.

Irons J.R., Weismiller R.A., Petersen G.W. (1989). Soil reflectance. *In: G. Asrar Ed., Theory and Applications of Optical Remote Sensing*. Wiley, New York, 66-106.

Islam K., Singh B., McBratney A. (2003). Simultaneous estimation of several soil properties by ultra-violet, visible, and near-infrared reflectance spectroscopy. *Australian Journal of Soil Research,* **41**, 1101-1114.

Jackson M.L. (1964). Análisis Químico de Suelos. Ed. Omega, Barcelona.

Janik L.J., Skjemstad J.O. (1995). Characterization and analysis of soils using mid infrared partial least-squares. 2. Correlations with some laboratory data. *Australian Journal of Soil Research*, **334**, 637-650.

Jørgensen A. (2000). Clustering excipient near infrared spectra using different chemometric methods. Technical report, Dept. of Pharmacy, University of Helsinki.

Karavanova E. I., Shrestha D.P., Orlov D.S. (2001). Application of remote sensing techniques for the study of soil salinity in semi-arid Uzbekistan. *In: Bridges E.M. et al. (Eds). Response to Land Degradation*. Oxford & IBH Publishing Co. Pvt. Ltd, 261-273.

Karimi Y., Prasher S.O, Madani A., Kim S. (2008). Application of support vector machine technology for the estimation of crop biophysical parameters using aerial hyperspectral observations. *Canadian Biosystems Engineering*, **50**, 713-720.

Kawano S. (2002). Sampling and sample presentation. *In: Siesler Hw, Ozaki Y, Kawata S, Heise HM (Ed). Near-infrared spectroscopy. Principles, instruments, applications*. Wiley-VCH Verlag, Weinheim, 115-124.

Kawata S. (2002). Instrumentation for Near-Infrared Spectroscopy. *In: Siesler Hw, Ozaki Y, Kawata S, Heise HM (Ed). Near-infrared spectroscopy. Principles, instruments, applications*.

Wiley-VCH Verlag, Weinheim, 43-73.

Khosla R., Inman D., Westfall D.G., Reich R. (2005). Spatial distribution of plant nitrogen uptake across site-specific management zones. Western Nutrient Management Conference (6), Salt Lake City, UT.

Kirshnan P., Alexander J.D., Butler B.J., Hummel J.W. (1980). Reflectance technique for predicting soil organic matter. S*oil Science Society of America Journal*, **44**, 1282–1285.

Kizos T., Koulouri M. (2005). Economy, demographic changes and morphological transformation of the agricultural landscape of Lesvos, Greece. *Human Ecology Review*, **12**, 183-192.

Knight F.W., Levinson S.H. (2012). Cuba. Encyclopaedia Britannica. http://www.britannica.com/EBchecked/topic/145542/Cuba/233416/Soils

Koch B., Khosla R. (2003). The role of precision agriculture in cropping systems. *Journal of Crop*, **8**, 361–381.

Kovačević M., Bajat B., Gajić B. (2010). Soil type classification and estimation of soil properties using support vector machines. *Geoderma*, **154** (3-4), 340–347.

Kovar J.L., Pierzynski G.M. (Eds) (2009). Methods of phosphorus analysis for soils, sediments, residuals, and waters. Southern Cooperative Series Bulletin 408, Southern Extension–Research Activity–17 (SERA 17). Available in: http://www.sera17.ext.vt.edu/Documents/P_Methods2ndEdition2009.pdf.

Kulmala A. (2012). Split application of nitrogen. Putting best agricultural practices into work. Available in: http://www.balticdeal.eu/measure/split-application-of-nitrogen-n/

Ladoni M., Bahrami H. A., Alavipanah S.K, Norouzi A.A. (2010). Estimating soil organic carbon from soil reflectance: a review. *Precision Agriculture*, **11**, 82–99.

Latz K., Weismiller R. A., Scoyoc G.E.V., Baumgardner M.F. (1984). Characteristic variations in spectral reflectance of selected eroded alfisols. *Soil Science Society of America Journal*, **48** (5), 1130-1134.

Leblon B. (1997). Soil and Vegetation Optical Properties. Faculty of Forestry and Environmental Management University of New Brunswick, Fredericton (NB), Canada.

Lee W.S., Sánchez J.F., Mylavarapu Y.R.S., Choe J.S. (2003). Estimating chemical properties of Florida soils using spectral reflectance. *Transactions of the ASAE*, **46** (5), 1443-1453.

Leung D.H.Y. (2005). Cross-validation in nonparametric regression with outliers. *The Annals of Statistics*, **33** (5), 2291-2310.

Lv G., Yang, H. (2011). Discrimination of different brands of Nescafé coffee using vis-NIR spectroscopy and comparative study. *Advances in Biochemical Engineering*, 1-2, 163–166.

Lobell D.B., Asner G.P. (2002). Moisture effects on soil reflectance. *Soil Science Society of America Journal*, **66** (3), 722.

López G., Fuentes E., Vázquez H. (1981). Resumen sobre los elementos fundamentales que deben ser redactados encada epígrafe del informe de Suelos por municipio a escala 1:25 000. Dpto. de Suelos y agroquímica. Ministerio de la Agricultura. Dir. Gral. de suelos y fertilizantes.

López-Granados F., Jurado-Expósito M., Peña-Barragán J.M., García-Torres L. (2005). Using geostatistical and remote sensing approaches for mapping soil properties. *European Journal of Agronomy*, **23**, 279-289.

Lv G., Yang H. (2011). Discrimination of different brands of Nescafé coffee using Vis/NIR spectroscopy and comparative study. *Advances in Biochemical Engineering*, **1-2**, 163-166.

McGraw-Hill (2003). Dictionary of Scientific & Technical Terms, 6th ed. The McGraw-Hill Companies, Inc.

Magdoff F., Lanyon L., Liebhardt B. (1997). Nutrient cycling, transformations, and flows: implications for a more sustainable agriculture. *Advances in Agronomy*, **60**, 2-68.

Mahinakbarzadeh M., Simkins S., Veneman P.L.M. (1991). Spatial variability of organic matter content in selected Massachussetts map units. *In*: *Mausbach M.J., Wilding L.P. (Eds). Proc. Int. Spatial variabilities of soils and landforms symp*. Las Vegas, Nevada, SSSA, Madison, WI., 231-242.

Maleki M.R., Van Holm L., Merckx R., Ramon H., De Baerdemaeker J., Mouazen A.M. (2006). Phosphorus sensing for fresh soils using visible and near infrared spectroscopy. *Biosystems Engineering*, **95** (3), 425-436.

Malley D.F., Yesmin L., Eilers R.G. (2002). Rapid analysis of hog manure and manure-amended soils using near-infrared spectroscopy. *Journal of the American Soil Science Society*, **66**, 1677-1686.

Malley D.F., Martin P.D., Ben–Dor E. (2004). Application in analysis of soils. *In: Roberts C.A. et al. (Eds). New-infrared spectroscopy in agriculture*. ASA, CSSA, SSSA, Madison, WI., 729-784.

Mandal S., Thakur T.C. (2010). Design and development of subsoiler-cum-differential rate fertilizer applicator. Agricultural Engineering International: *CIGR Journal*, **12**(1), 74 - 83.

Mann K.K., Schumann A.W., Obreza T.A., Teplitski M., Harris W.G., Sartain J.B. (2011). Spatial Variability of Soil Chemical and Biological Properties in Florida Citrus Production. *Soil Science Society of America Journal*, **75**, 1863–1873.

McCarty G.W., Reeves J. B., Reeves V. B., Follett R. F., Kimble J.M. (2002). Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. *Soil Science Society of America Journal*, **662**, 640-646.

Meteorological Centre from Villa Clara province (2011). Marta Abreu, 57, Santa Clara, Villa Clara, Cuba.

Meyer J. H., Van Vuuren J. A. J., Van Antwerpen R. (2005). The potential role of Near Infrared Reflectance (NIR) monitoring in precision agriculture (PAG). *Proc S Afr Sug Technol Ass*, **79**, 202.

Minasny B., McBratney A.B., Bellon-Maurel V., Roger J.M., Gobrecht A., Ferrand L., Joalland S. (2011). Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon. *Geoderma,* (167-168), 118-124.

Monsef Z. (2012). Reflectance IR Spectroscopy, Infrared Spectroscopy - Materials Science, Engineering and Technology, Prof. Theophanides Theophile (Ed), ISBN: 978-953-51-0537-4, InTech. Available in: http://www.intechopen.com/books/infrared-spectroscopy-materials-science-engineering-andtechnology/fundamental-of-reflectance-ir-spectroscopy

Montgomery O.L. (1976). An investigation of the relationship between spectral reflectance and the chemical, physical and genetic characteristics of soils, PhD thesis, Purdue University.

Moore I.D., Gessler P.E., Nielsen G.A., Peterson G.A. (1993). Soil attribute prediction using terrain analysis. *Soil Science Society of America Journal*, **57**, 443-452.

Moran M.S., Inoue Y., Barnes E.M. (1997). Opportunities and limitations for image-based remote sensing in precision crop management. *Remote Sensing of Environment*, **61** (3), 319-346.

Morón A., Cozzolino D. (2007). Measurement of phosphorus in soils by near infrared reflectance spectroscopy: Effect of reference method on calibration. *Communications in Soil Science and Plant Analysis*, **38** (15-16), 1965-1974.

Mouazen A.M., De Baerdemaeker J., Ramon H. (2005). Towards development of on-line soil moisture content sensor using a fibre-type NIR spectrophotometer. *Soil and Tillage Research*, **80**, 171-183.

Mouazen A.M., Karoui R., De Baerdemaeker J., Ramon H. (2006). Characterization of soil water content using measured visible and near infrared spectra. *Soil Science Society of America Journal*, **70**, 1295-1302.

Mouazen A.M., Maleki M.R., De Baerdemaeker J., Ramon H. (2007). On-line measurement of some selected soil properties using a VIS–NIR sensor. *Soil and Tillage Research*, **93** (1), 13-27.

Mouazen A.M. (2011). Field vegetables: on-line measurement of selected soil properties towards the refinement of fertilisation management – HDC studenship. Agriculture and Horticulture Development Board.

Müller A.L.H., Flores É.M.M., Müller E.I., Silva F.E.B., Ferrão M.F. (2011). Attenuated Total Reflectance with Fourier Transform Infrared Spectroscopy (ATR/FTIR) and Different PLS Algorithms for Simultaneous Determination of Clavulanic Acid and Amoxicillin in Powder Pharmaceutical Formulation. *Journal of the Brazilian Chemical Society*, **22** (10), 1903-1912.

Murphy C.L., Macleod A.P., Chapman G.A., Milford H.B., McGaw A.J. E., Edye J.A., Simons N.A. (2001). NSW State Soil Landscape Mapping Program and Derivative Products. *In: Proceedings of the Geospatial Information and Agriculture Symposium*, Sydney.

Mutuo P.K., Shepherd K.D., Albrecht A., Cadisch G. (2006). Prediction of carbon mineralization rates from different soil physical fractions using diffuse reflectance spectroscopy. *Soil Biology & Biochemistry*, **38**, 1658-1664.

Mzuku M., Khosla R., Reich R., Inman D., Smith F., MacDonald L. (2005). Spatial variability of measured soil properties across site-specific management zones. *Soil Science Society of America Journal*, **69**, 1 572-1 579.

Naes T., Isakson T., Fearn T., Davies T. (2002). A User-Friendly Guide to Multivariate Calibration and Classification. NIR Publications, Chichester, UK.

Nduwamungu C., Ziadi N., Tremblay G.F., Parent L.E. (2009). Near-infrared reflectance spectroscopy prediction of soil properties: Effects of sample cups and preparation. *Soil Science Society of America Journal*, **73**, 1896-1903.

Nezat C.A., Blum J.D., Yanai R.D., Park B.B. (2008). Mineral sources of calcium and phosphorus in soils of the northeastern United States. *Soil Science Society of America Journal*, **72**, 1786–1794.

Nørgaard L., Saudland A., Wagner J., Nielsen J.P., Munck L., Engelsen S.B. (2000). Interval Partial Least-Squares Regression (iPLS): A Comparative Chemometric Study with an Example from Near-Infrared Spectroscopy. *Applied Spectroscopy*, **54** (3), 413-419.

Norris H.K., Davies T. (2012). Examining diffuse reflection and transmission spectra more thoroughly: Part 2. Sample temperature and sample composition. *Spectroscopy Asia*, **2** (8), 12-17.

Nowkandeh S.M., Homaee M., Noroozi A.A. (2013). Mapping Soil Organic Matter Using Hyperion Images. *International Journal of Agronomy and Plant Production*, **4** (8), 1753-1759.

Oliveira Sa S., Ferrao M.F., Galdos M.V., Machado Bittar C.M., Poppi R.J. (2010). Application of LS-SVM-NIR spectroscopy for carbon and nitrogen prediction in soils under sugarcane. *19th World Congress of Soil Science, Soil Solutions for a Changing World,* 1- 6 August, Brisbane, Australia. Published on DVD.

Olsen S.R., Sommers L.E. (1982). Phosphorus. *In*: *Pageetal A.L. (Ed). Methods of soil analysis*, *part 2.* Agron. Mongr. 9. 2nd ed. ASA, SSSA, Madison, WI., 403-430.

ONE (2009). Villa Clara. Desarrollo económico social en el período revolucionario. Oficina Nacional de Estadísticas, Santa Clara, Villa Clara. Available in: http://www.vcl.jovenclub.cu/prov/onevc/images/PDF/desarrollo.pdf

ONEI (2010). 26. Villa Clara. Oficina Nacional de Estadísticas e Información. Available in: http://www.one.cu/publicaciones/provincias_masinf/villa%20clara.htm

Oniani O.G., Chater M., Mattingly G.E.G. (1973). Some effects of fertilisers and farmyard manure on the organic phosphorus in soils. *Journal of Soil Science*, **24**, 1–9.

Optical Society of America (2008). The reflection of light. Available in: http://www.optics4kids.org/home/teachersparents/articles/the-reflection-of-light/

Owen T. (1996). Fundamentals of UV-visible spectroscopy: a primer. Hewlett-Packard Company, Germany, 61.

Parkin T.B. (1993). Spatial variability of microbial processes in soil: a review. *Journal of Environmental Quality,* **22**, 409-417.

Patil S.S., Patil V.C., Al-Gaadi K.A. (2011). Spatial variability in fertility status of surface soils. *World Applied Sciences Journal*, **14**, 1020-1024.

Pérez H., Rodríguez I., Betancourt Y., Gómez J.R., Más R., Rodríguez L. (2001). Informe. Evaluación de la Aptitud Física de las Tierras del CAI. Primera Aproximación. Instituto Nacional de Investigaciones de la Caña de Azúcar, La Habana, 67.

Peters P. (2003). Cutting losses: Cuba down sizes its sugar industry. Lexington Institute, 14.

Phillips S. (2009). The role of spatial variability in nutrient management, Fall 9. International Plant Nutrition Institute (IPNI), Georgia, USA.

Pleysier J.L. (1995). Soil Sampling and Sample Preparation. Uitgever International Inst. of Tropical Agriculture, (2) van IITA research guide, International Institute of Tropical Agriculture Ibadan.

Prasad R., Pandey A., Jha S.K., Singh K P., Yadav G.S. (2011). Classification of fields having different soil moisture content by SVM technique using bistatic scatterometer. *Recent Research in Science and Technology*, **3** (1), 105-113.

Ramirez-Lopez L., Behrens T., Schmidt K., Stevens A., Demattê J.A.M., Scholten T. (2013). The spectrum-based learner: A new local approach for modeling soil vis–NIR spectra of complex datasets. *Geoderma*, **195-196**, 268-279.

Reeves J.B., Francis B.A., Hamilton S.K. (2005). Specular reflection and diffuse reflectance spectroscopy of soils. *Applied Spectroscopy*, **59**, 39-46.

Rehbein H., Oehlenschlager J. (2009). Fishery products: Quality, safety, authenticity. Wiley-Blackwell, UK.

Robert P.C. (2002) Precision agriculture: a challenge for crop nutrition management. *Plant and Soil,* **1** (247), 143-149.

Robinson N.J., Rampant P.C., Callinan A.P.L., Rab M.A., Fisher P.D. (2009). Advances in precision agriculture in south-eastern Australia. II. Spatio-temporal prediction of crop yield using terrain derivatives and proximally sensed data. *Crop and Pasture Science*, **60**, 859-869.

Roger J.M., Chauchard F., Bellon-Maurel V. (2003). EPO-PLS external parameter orthogonalisation of PLS application to temperature-independent measurement of sugar content of intact fruits. *Chemometrics and Intelligent Laboratory Systems,* **66**, 191–204.

Rowse H.R., Stone D.A. (1980). Deep cultivation of a sandy loam, I. Effects on growth, yield and nutrient content of potatoes, broad beans, summer cabbage and red beet in 1977. *Soil and Tillage Research*, **1**(1), 57-68.

Saeys W., Ramon H. (2004). Rapid analysis of hog manure using visual and near-infrared reflectance spectroscopy. *In: AgEng 2004 Conference "Engineering the future location"*, Leuven, Belgium, 12-16 September.

Sahrawat K.L. (1982). Simple modification of the Walkley-Black method for simultaneous determination of organic carbon and potentially mineralizable nitrogen in tropical rice soils. *Plant and Soil*, **69**, 73-77.

Sauer M., Hofkens J., Enderlein J. (2008). Handbook of fluorescence spectroscopy and imaging: From ensemble to single molecules. Weinheim: Wiley-VCH.

Schirrmann M., Domsch H. (2011). Sampling procedure simulating on-the-go sensing for soil nutrients. *Journal of Plant Nutrition and Soil Science*, **174**, 333-343.

Schirrmann M., Gebbers R., Kramer E., Seidel J. (2011). Soil pH-mapping with an on-the-go sensor. *Sensors*, **11**, 573-598.

Schirrmann M., Kramer E., Gebbers R. (2012). Field scale mapping of soil fertility parameters by combination of proximal soil sensors. *In: International Conference of Agricultural Engineering CIGR AgEng 2012*. Diazotec, S. L., Valencia, (ISSN 84-615-9654-4), p. 1-6.

Schubert W.J. (1965). Lignin Biochemistry. Academic Press, New York, 474.

Schumacher B.A. (2002). Methods for the determination of total organic carbon (TOC) in soils and sediments. United States Environmental Protection Agency, Ecological Risk Assesment Support Center.

Schwartz G., Eshel G., Ben-Dor E. (2011). Reflectance Spectroscopy as a Tool for Monitoring Contaminated Soils. *In: Pascucci S. (Ed). Soil Contamination*. InTech, Rijeka, Croatia, 70-71.

Scott J.T., Haggard B.E., Sharpley A.N., Romeis J.J. (2011). Change point analysis of phosphorus trends in the Illinois River (Oklahoma) demonstrates the effects of watershed management. *Journal of Environmental Quality,* **40** (4), 1249-1256.

Shao Y., He Y. (2011). Nitrogen, phosphorus, and potassium prediction in soils, using infrared spectroscopy. *Soil Research*, **49**, 166-172.

Sharma B.K. (2005). Water pollution. GOEL Publishing House, Meerut.

Shifteh S.B., Hassanpour F., Ezani A., Miremadi S.R., Tabari H. (2011). Investigation of spatial variability and pattern analysis of soil properties in the northwest of Iran. *Environmental Earth Sciences*, **64**, 1849-1864.

Shimadzu (2012). ABC's of the Diffuse Reflection Method. Analytical and Measuring Instruments Available in: http://www.shimadzu.com/an/ftir/support/ftirtalk/talk1/intro.html

Siesler H.W. (2002). Introduction. *In: Siesler H.W., Ozaki Y., Kawata S., Heise H.M. (Ed). Near-infrared spectroscopy. Principles, instruments, applications*. Wiley-VCH Verlag, Weinheim, 1-10.

Sinfield J.V., Fagerman D., Colic O. (2010). Evaluation of sensing technologies for on-the-go detection of macro-nutrients in cultivated soils. *Computers and Electronics in Agriculture*, **70**, 1-18.

Six J., Bossuyt H, Degryze S., Denef K. (2004). A history of research on the link between micro-aggregates, soil biota, and soil organic matter dynamics. *Soil and Tillage Research*, **79**, 7–31.

Six L., Smolders E., Merckx R. (2013). The performance of DGT versus conventional soil phosphorus tests in tropical soils-maize and rice responses to P application. *Plant and Soil,* **1-2** (366), 49–66.

Smola A.J., Scholkopf B. (1998). A tutorial on support vector regression. Produced as part of the ESPRIT Working Group in Neural and Computational Learning II. NeuroCOLT2 27150.

Sörensen L.K., Dalsgaard S. (2005). Determination of clay and other soil properties by near infrared spectroscopy. *Soil Science Society of America Journal*, **69**, 159–167.

Stenberg B., Viscarra Rossel R.A., Mouazen A.M., Wetterlind J. (2010). Visible and near infrared spectroscopy in soil science. *Advances in Agronomy*, **107**, 163-215.

Stevens A., Van Wesemael B., Bartholomeus H., Rosillon D., Tychon B., Ben-Dor E. (2008). Laboratory, field and airborne spectroscopy for monitoring organic carbon content in agricultural soils. *Geoderma*, ***144***, 395-404.

Stuart B.H. (2004). Infrared spectroscopy: fundamentals and 450 applications, John Wiley & Sons Ltd, Chichester, UK, 244.

Sun H., Li M., Zhao Y., Li X., Liang J., Chen L. (2009). Characteristics of soil spectral reflectance and estimation of soil parameters in Fuxin opencast coal mine. *Spectroscopy and Spectral Analysis*, **29** (12), 3365-3368.

Suykens J.A.K., Van Gestel T., De Brabanter J., De Moor B., Vandewalle J. (2002). *Least Squares Support Vector Machines*, World Scientific, Singapore.

Tekin Y., Kuang B., Mouazen A.M. (2013). Potential of On-Line Visible and Near Infrared Spectroscopy for Measurement of pH for Deriving Variable Rate Lime Recommendations. *Sensors*, **13**, 10177-10190.

Thomasson J.A., Sui R., Cox M.S., Al–Rajehy A. (2001). Soil reflectance sensing for determining soil properties in precision agriculture. *Transactions of the ASAE*, **44**(6), 1445–1453.

Thomsen I.K., Bruun S., Jensen L.S., Christensen B.T. (2009). Assessing soil carbon lability by near infrared spectroscopy and NaOCl oxidation. *Soil Biology & Biochemistry*, **10** (41), 2170-2177.

Thrikawala S., Weersink A., Kachanoski G., Fox G. (1999). Economic feasibility of variable-rate technology for nitrogen on corn. *American journal of agricultural economics*, **81**, 914-927.

Twarakavi N.K.C., Simunek J., Schaap M.G. (2009). Development of pedotransfer functions for estimation of soil hydraulic parameters using support vector machines. *Soil Science Society of America Journal*, **73**, 1443-1452.

Ulmschneider M., Roggo Y. (2008). Process analytical technology, *In: Gad S.C. (Ed), Pharmaceutical Manufacturing Handbook*. Hoboken, N.J., Wiley-Interscience, 353-410.

UMN (2002). Why Manage Soil Organic Matter? Regents of the University of Minnesota. Available in: http://www.extension.umn.edu/distribution/cropsystems/components/7402_02.html

Valentine K.W.G., Sprout P.N., Baker T.E., Lawkulich L.M. (Eds.) (1978). The Soil Landscapes of British Columbia. BC Ministry of Environment, Resource Analysis Branch, 197. Available in: http://www.env.gov.bc.ca/soils/landscape/index.html

van Schöll L., Nieuwenhuis R. (2004). Soil fertility management. 4th ed. Agromisa Foundation, Wageningen, 83.

Vapnik V. (1995). The Nature of Statistical Learning Theory. New York, NY: Springer-Verlag.

Vapnik V., Golowich S., Smola A. (1997). Support vector method for function approximation, regression estimation, and signal processing. *In: Advances in Neural Information Processing Systems 9*; Mozer M., Jordan M., Petsche T. (Eds). MIT Press: Cambridge, MA. Pp: 281-287.

Vasques G.M., Grunwald S., Sickman J. O. (2009). Modeling of Soil Organic Carbon Fractions Using Visible-Near-Infrared Spectroscopy. *Soil Science Society of America journal*, **1** (73), 176-184.

Veris (2012). VIS-NIR - Spectrophotometer: Addressing tomorrow's soil challenges today. Available in: http://www.veristech.com/products/visnir.aspx

Veseth R. (1986). Managing Variable Soils PNW Conservation tillage handbook series. Chapter 6 - Fertility (6), Fall.

Villegas R., Ponce de León D., Balmaseda C., Pineda E. (1995). Cuba: Brown Calcareous Reference Soils derived from limestone. *Soil Brief Cuba 5*. Instituto Nacional de Investigaciones de la caña de azúcar, Habana and International Soil Reference and Information Center, Wageningen, 26.

Villegas R., De León M., Campos J., Valdés A., Fernández A., Menéndez A., Cabrera A., Pérez M. (1999). SERFE. Manual de procedimiento. Ministerio del azúcar. Instituto Nacional de Investigaciones de la caña de Azúcar. Servicio de Recomendaciones de fertilizantes y enmiendas.

Viscarra Rossel R.A., Walvoort D.J.J., McBratney A.B., Janik L.J., Skjemstad J.O. (2006). Visible, near-infrared, mid-infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, **131**, 59-75.

Viscarra Rossel R.A., Taylor H.J., McBratney A.B. (2007) Multivariate calibration of hyperspectral γ-ray energy spectra for proximal soil sensing. *European Journal of Soil Science*, **58**, 343–353.

Viscarra Rossel R.A., Fouad Y., Walter C. (2008) Using a digital camera to measure soil organic carbon and iron contents. *Biosystems Engineering*, **100**, 149-159.

Viscarra Rossel R.A., Behrens T. (2010). Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma*, **158**, 46-54.

Vo K. (2010). Spectrophotometry. The Regents of the University of California, Davis. Available in: http://chemwiki.ucdavis.edu/@api/deki/pages/345/pdf

Volkan A., van Es H.M., Akbas F., Durak A., Hively W.D. (2010). Visible near infrared reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey. *Journal of Arid Environments*, **74**, 229-238.

Võsa T., Nugis E., Vennik K., Meripõld H., Viil P., Kuht J. (2009). Some possibilities of studying the precision farming in Estonia, methods and results of complex investigation *Computer and Computing Technologies in Agriculture,* **2** (1), 1-7.

Walkley A., Black I.A. (1934). An examination of Degtjareff method for determining soil organic matter and a proposed modification of the chromic acid titration method. *Soil Science,* **37**, 29-37.

Wang S.Q., Shu N., Zhang H.T. (2008). In-site total N content prediction of soil with Vis/NIR spectroscopy. *Guang Pu Xue Yu Guang Pu Fen Xi*, **28** (4), 808-12.

Wang J., Qin Q., Dong H., Chen Ch., Meng Q. (2013). Study on quantitative retrieval of soil organic matter based on bare soil spectrum. *In: 3rd International Conference on Multimedia Technology (ICMT 2013)*.

Warwick H. (2000). Cuba's organic revolution. Third World Resurgence, 118/119.

Weindorf D.C., Zhu Y. (2010). Spatial variability of soil properties at Capulin Volcano, New Mexico, USA: implications for sampling strategy. *Pedosphere*, **20**, 185-97.

Wetterlind J., Stenberg B., Jonsson A. (2008). Near infrared reflectance spectroscopy compared with soil clay and organic matter content for estimating within-field variation in N uptake in cereals. *Plant Soil*, **302**, 317-327.

Whitbread A.M. (1995). Soil organic matter: its fractionation and role in soil structure. *Soil Organic Matter Management for Sustainable Agriculture. In: Lefroy R.D.B., Blair G.J., Craswell E.T. (Eds). ACIAR Proceedings, 56*, 124-30.

Workman J. Jr., Weyer L. (2008). Practical Guide to Interpretive Near-Infrared Spectroscopy. Taylor & Francis Group: Boca Raton, FL, USA, 332.

Wülfert F. (2004) Temperature-robust multivariate calibration, PhD. Thesis, 135. University of Amsterdam, The Netherlands. Available in: http://dare.uva.nl/record/161626

Yang H., Kuang B., Mouazen A.M. (2011). Prediction of soil TN and TC at a farm-scale using VIS-NIR spectroscopy. *Advanced Materials Research*, **225-226** (1-2), 1258-1261.

Yang H., Mouazen A.M. (2012). Vis/near andmid-infrared spectroscopy for predicting soil N and C at a farmscale. *In: Theophanides T. (Ed). Infrared Spectroscopy-Life and Biomedical Sciences*, InTech Press, Rijeka, Croatia, 185–210.

Yu S., Yao Y., Wang D., Si H. (2013). Studies on soil organic matter content mapping using EO-1 hyperion data. *In: Second International Conference on Agro-Geoinformatics 2013,* 439 – 443.

Zaman Q.U., Schumann A.W., Miller W.M. (2005). Variable rate nitrogen application in Florida citrus based on ultrasonically-sensed tree size. *Applied Engineering in Agriculture*, 21, 331–335.

Zhang H., Raun B. (2006). *Oklahoma soil fertility handbook*. Department of Plant and Soil Sciences, Oklahoma State University, 150.

Zhang H.H., Liu Y., Wu Y., Zhu J. (2008). Variable selection for multicategory SVM via sup-norm regularization. *Electronic Journal of Statistics*, **2**, 149-167.